

电子固废处理析取过程溶解氧浓度预测模型

韩红桂^{1, 2, 3} 任晨昊^{2, 3} 李方昱^{1, 3} 刘 峥^{1, 3}

摘 要 电子固废处理析取过程涉及多变量强耦合的多相反应, 其原料异质性使工艺变量间呈现动态因果特性, 导致表征反应平衡的溶解氧浓度难以准确预测. 为解决该问题, 设计一种电子固废处理析取过程溶解氧浓度自适应因果推理预测模型. 首先, 提出一种基于多头注意力的自适应因果发现方法, 更新工艺变量的因果注意力分布, 实现动态因果结构的建模. 其次, 设计一种基于动态掩码的因果干预策略, 抑制工艺参数的虚假因果相关, 实现模型真实因果关系重构. 最后, 构建一种基于元更新的门控网络结构, 调节时空特征与因果特征的门控融合权重, 实现析取过程溶解氧浓度的稳定预测. 在电子固废处理析取过程数据集上验证该预测模型, 实验结果表明, 该模型能够准确描述工艺变量间动态因果关系, 提升溶解氧浓度的预测精度.

关键词 电子固废处理; 析取过程; 多头注意力; 因果干预; 元更新

引用格式 韩红桂, 任晨昊, 李方昱, 刘峥. 电子固废处理析取过程溶解氧浓度预测模型. 自动化学报, 2026, **52**(8): 1781–1790

DOI 10.16383/j.aas.c250658 **CSTR** 32138.14.j.aas.c250658

Prediction Model of Dissolved Oxygen Concentration in Electronic Waste Treatment Disjunction Process

HAN Hong-Gui^{1, 2, 3} REN Chen-Hao^{2, 3} LI Fang-Yu^{1, 3} LIU Zheng^{1, 3}

Abstract The electronic waste treatment disjunction process involves multiphase reactions with multivariable strong coupling. The heterogeneity of raw materials results in dynamic causal characteristics in process variables, complicating the accurate prediction of dissolved oxygen concentration at reaction equilibrium. To address the issue, an adaptive causal inference prediction model is designed for the dissolved oxygen concentration in electronic waste treatment disjunction process. First, an adaptive causal discovery method based on multi-head attention is proposed, which updates the causal attention distribution among process variables to model dynamic causal structure. Second, a causal intervention strategy based on the dynamic mask is established to suppress spurious causal correlations among process parameters, enabling the reconstruction of true causal relationships in the model. Finally, a gating network structure based on meta-update is constructed to adjust the gating fusion weights of spatiotemporal and causal features, achieving stable prediction of dissolved oxygen concentration in the disjunction process. The proposed prediction model is validated on the electronic waste treatment disjunction process dataset. Experimental results demonstrated that the model can accurately characterize dynamic causal relationships among process variables and improve the prediction accuracy of dissolved oxygen concentration.

Keywords electronic waste treatment; disjunction process; multi-head attention; causal intervention; meta-update

Citation Han Hong-Gui, Ren Chen-Hao, Li Fang-Yu, Liu Zheng. Prediction model of dissolved oxygen concentration in electronic waste treatment disjunction process. *Acta Automatica Sinica*, 2026, **52**(8): 1781–1790

收稿日期 2025-11-18 录用日期 2026-03-21
Manuscript received November 18, 2025; accepted March 21, 2026

2030 新一代人工智能重大专项项目 (2025ZD0123700), 国家自然科学基金 (62125301, 62021003, 62373014, 92467205), 北京市科技新星计划 (20240484694), 青年北京学者计划 (037) 资助

Supported by 2030 New Generation Artificial Intelligence National Science and Technology Major Project (2025ZD0123700), National Natural Science Foundation of China (62125301, 62021003, 62373014, 92467205), Beijing Nova Program (20240484694), and Beijing Youth Scholar (037)

本文责任编辑 孟凡利

Recommended by Associate Editor MENG Fan-Li

1. 北京工业大学信息科学技术学院 北京 100124 2. 北京工业大学环境科学与工程学院 北京 100124 3. 数字社区教育部工程研究中心 北京 100124

1. School of Information Science and Technology, Beijing University of Technology, Beijing 100124 2. College of Environmental Science and Engineering, Beijing University of Technology, Beijing 100124 3. Engineering Research Center of Digital Community, Ministry of Education, Beijing 100124

电子产品的快速更新迭代加剧了电子固废堆积, 未经处理的电子固废大量堆积导致严重的环境污染^[1–3]. 在电子固废处理过程中, 从拆解材料中提取稀贵金属的过程被称为电子固废处理析取过程 (electronic waste treatment disjunction process, EWTDP). EWTDP 涉及电子产品的各种拆解材料, 例如含金属的外壳、电路板和电池. 反应过程包括非金属物质的去除, 以及稀有金属和贵金属的分离和提纯; 同时涉及一系列复杂的理化反应, 需要串联和并联的多个反应器协调运行, 最终实现拆解

logy, Beijing 100124 3. Engineering Research Center of Digital Community, Ministry of Education, Beijing 100124

材料的高效转化、分离和提纯。多元时间序列预测在析取过程中具有重要的应用价值^[4-5],通过揭示过程变量之间的复杂时空相关性,为反应过程关键变量预测提供技术支撑^[6-7]。电子固废的成分具有高度复杂性与动态变化特征,不同来源的电子固废在金属含量、氧化程度及杂质比例上存在显著差异,析取反应机理也呈现高度非线性特征。准确建模多变量动态关系是电子固废处理析取过程中亟待解决的问题。

EWTDP 的具体工艺过程及主要设备如图 1 所示。电子固废拆解材料、酸液和氧气等多相物料进行耦合反应,其动力学机制复杂,多工艺变量间的因果结构不明晰,受原料成分波动、操作条件变化以及多反应器协同运行等多重因素的影响。其中,溶解氧浓度直接决定稀有金属和贵金属的析出效率与最终产物的纯度。传统的机理模型往往依赖于简化的动力学假设,难以全面建立动态环境下变量之间的非线性关系;数据驱动方法虽然能够从历史数据中学习时序模式,但是在捕捉复杂的时空依赖和动态因果关联时仍存在不足。因此,需要开发一种能够同时建模时空依赖与因果动态特征的方法,以提升电子固废处理析取过程溶解氧浓度的预测能力。

EWTDP 涉及多变量动态关联和复杂的时空依赖,因此,时空图神经网络在该应用场景下具有显著优势^[8-9]。Qi 等^[10]提出一种结合注意力机制的时空图神经网络模型,通过融合图注意力网络、双向门控循环单元,解决了时间序列预测中时空特征提取不充分的问题。Xu 等^[11]提出一种多注意力时空融合图神经网络,通过动态稀疏图构建、双通道

时空特征融合及多级注意力机制,解决了数据驱动方法在复杂工业过程中难以有效建模非线性时空关联的问题。Yu 等^[12]提出一种时空图卷积网络,通过将图卷积与门控时序卷积结合,解决了传统预测方法难以有效捕捉复杂时空依赖性的问题。上述方法有效提升了模型对复杂时空关系的建模能力。然而,现有时空图神经网络依赖于静态或预定义图结构,难以适应析取过程中由原料波动、工况迁移引发的时空依赖变化问题。

为解决时空图神经网络对动态时空相关性建模能力不足的问题^[13-14],Sun 等^[15]提出一种自适应邻域选择时空图神经网络,通过自适应邻域选择机制动态优化图结构,解决了固定图结构限制时空关系建模的问题,提升了时空图神经网络的泛化性能。Zheng 等^[16]提出一种关联感知的时空图学习方法,通过动态图学习模块和时空图网络整合多跳空间依赖与多尺度时间模式,解决了多元时序异常检测中动态相关性建模不足的问题,显著提升了检测精度。Li 等^[17]提出一种动态时空图神经网络,通过动态图估计和自适应引导传播,解决了现有图神经网络无法捕捉动态依赖的问题,提升了模型的预测性能。然而,在电子固废处理析取过程中,仅基于统计相关性建模的方法易产生变量间的虚假依赖,导致对析取反应机理的误判,影响系统的稳定运行。

为精准构建变量间的因果约束关系^[18-19],Miraki 等^[20]提出一种可解释因果图神经网络,融合因果图发现、时间卷积和门控图神经网络,提升了预测可解释性和准确性。Wang 等^[21]提出一种因果图注意力网络,通过因果分析和反向调整策略,提高了模型的准确性和稳定性。Luo 等^[22]提出一种因果干预

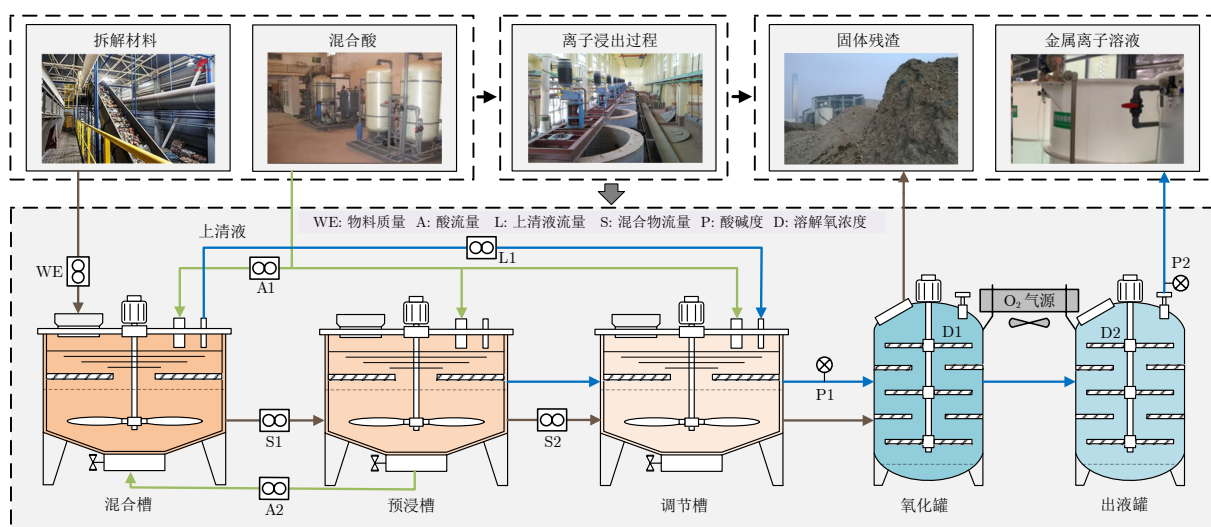


图 1 电子固废处理析取过程示意图

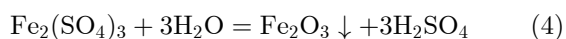
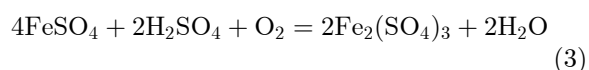
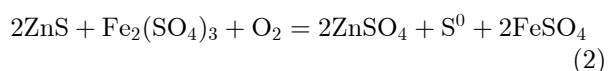
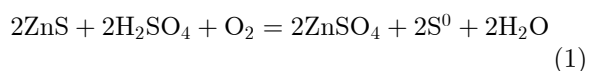
Fig.1 Schematic diagram of electronic waste treatment disjunction process

时空图神经网络, 通过因果推理和随机干预机制, 解决了时空数据中隐藏混杂因素导致的虚假关联问题. 上述建模方法通过结合因果推理提高了时空图神经网络的预测精度, 但建模析取过程数据高非线性与动态因果特性是亟待解决的难题.

综上所述, 本文提出一种电子固废处理析取过程溶解氧浓度自适应因果推理预测模型 (adaptive causal inference prediction model, ACIPM). 该模型的主要贡献如下: 1) 提出一种多头注意力引导的自适应因果发现方法, 有效提升了模型对工艺变量间动态因果关系的提取能力, 精准建立动态因果结构; 2) 设计一种动态掩码操作的因果干预策略, 抑制了由非因果特征引起的虚假相关, 增强了析取过程模型的可解释性, 减少了无效信息的传播; 3) 构建元更新门控网络结构, 通过动态调节因果特征与时空特征的融合权重, 显著提高了模型对析取过程溶解氧浓度的预测精度.

1 电子固废处理析取过程运行特性分析

电子固废处理析取过程是一个高度复杂的工艺, 涉及多种物理与化学反应, 包括非金属组分溶解、有价金属分离与浸出以及离子提纯等关键步骤. 锌被广泛应用于镀锌、黄铜铸造和电池制造等领域, 在电子固废中通常以硫化物或氧化物的形式嵌于复杂的废物基体中, 若处置不当, 将导致严重的环境风险与资源浪费. 基于氧压酸浸的析取工艺可有效回收锌资源, 具体反应如下:



其中, ZnS 为电子固废拆解材料中的硫化锌; H_2SO_4 为硫酸溶液; O_2 为反应中的溶解氧; ZnSO_4 为硫酸锌溶液; S^0 为析出的硫单质; H_2O 为反应生成的水; $\text{Fe}_2(\text{SO}_4)_3$ 为铁氧化物与硫酸反应生成的硫酸铁; FeSO_4 为硫酸铁被还原生成硫酸亚铁; Fe_2O_3 为硫酸铁水解生成三氧化二铁沉淀; ZnO 为氧化锌.

在适当的溶解氧浓度及酸性环境下, ZnS 与 H_2SO_4 反应生成 ZnSO_4 , 并析出 S^0 . 该反应避免了生成二氧化硫气体; 电子固废成分复杂, 其中含有

大量铁氧化物, 与 H_2SO_4 反应后产生的 $\text{Fe}_2(\text{SO}_4)_3$ 可以作为氧化剂, 将 ZnS 中的硫氧化为 S^0 , 同时自身被还原为 FeSO_4 ; 在酸性介质中, 通入 O_2 将 FeSO_4 重新氧化为 $\text{Fe}_2(\text{SO}_4)_3$, 实现铁离子的循环使用, 降低试剂消耗; $\text{Fe}_2(\text{SO}_4)_3$ 水解生成 Fe_2O_3 , 并再生 H_2SO_4 , 实现废酸的回收以及铁氧化物的去除, 提高锌的回收效率; ZnO 与 H_2SO_4 反应生成 ZnSO_4 .

溶解氧浓度会直接影响硫化锌的氧化过程, 引发金属氧化动力不足或副反应, 导致杂质离子的剩余, 降低目标金属的选择性浸出效率, 同时造成污染气体的产生. 由于析取过程高放热、强腐蚀, 高精密度检测设备极易受到损坏, 设备维护成本较高, 导致影响工艺运行的关键变量——溶解氧浓度难以直接检测. 在工艺运行过程中, 电子固废由于型号的不同, 其组分会存在显著波动, 其中的金属含量、塑料比例和其他杂质的变化直接影响反应速率和产物纯度. 由于原料的异质性, 每批废弃物的拆解过程和所需化学添加剂的种类和用量不同. 此外, 在反应过程中需要使用专门的溶解剂和沉淀剂来分离不同金属, 而这些添加剂的浓度、投加时机及与 pH 条件的匹配均受到溶解氧的影响, 其浓度需精确预测以确保目标金属的最大回收率和最高纯度. 因此, 建立溶解氧浓度的预测模型是实现析取过程高效运行的重要挑战.

2 问题描述

在电子固废处理析取过程中, 溶解氧浓度受到多种动态耦合的工艺变量影响. 由于原料具有高度异质性, 工艺变量之间呈现出显著的动态因果特性, 其相关性会随时间不断波动. 同时, 该过程具有多时间尺度特征, 溶解氧浓度、pH 值等变量采样间隔为 1 min, 而物料成分、质量等变量采样间隔为 5 min. 然而, 大多数现有的时空图神经网络方法仅依赖由先验知识定义的静态图结构, 难以反映析取过程中变量关系在不同时间尺度上的动态变化. 此外, 在设备性能下降和测量噪声等外部扰动影响下, 仅基于统计相关性建模容易引入由隐藏混淆因素导致的虚假依赖, 使部分非真实因果关系参与特征传播, 进而造成预测偏差并降低模型的可解释性. 因此, 如何准确构建溶解氧浓度的动态因果模型, 是实现电子固废处理析取过程稳定运行亟须解决的关键难题. 为解决上述问题, 需构建一种因果时空建模方法, 该方法在时间尺度上与析取过程的溶解氧浓度采样频率保持一致, 通过在连续时间步上自适应更新, 捕捉快速工况扰动下的因果关系变化, 从而实现析取过程自适应因果相关性建模.

电子固废处理析取过程数据中的每个变量都被视为图结构中的一个节点,与每个节点相关的观测值被解析为节点特征或图信号.为了表征节点间的连接关系,采用图邻接矩阵进行形式化表达.使用 $G = (V, E)$ 表示变量间的动态耦合关系.其中,图结构包含:节点集 $V \in \mathbf{R}^N$;边集 E ,每条边连接两个顶点,表征变量间的特定关联;邻接矩阵 $A \in \mathbf{R}^{N \times N}$,描述节点间的连接强度,矩阵的维度为 $N \times N$, N 表示图中节点的数量.预测目标为:学习一个动态因果关系函数 $\mathcal{F}$,预测 $t+k$ 时刻所有节点的状态 $\mathbf{X}$.该过程可形式化表示为:

$$\mathbf{X}(t+k) = \mathcal{F}(\mathbf{X}(t), \mathbf{A}(t), \mathbf{C}(t)) \quad (6)$$

其中, $\mathcal{F}$ 表示整体预测函数; $\mathbf{X}(t)$ 为 t 时刻所有节点特征; $\mathbf{A}(t)$ 为节点间的因果关系强度; $\mathbf{C}(t)$ 为因果干预策略.

3 自适应因果推理预测模型

自适应因果推理预测模型分为以下三部分:基于多头注意力的自适应因果发现方法 (adaptive causal discovery method based on multi-head attention, ACDM)、基于动态掩码的因果干预策略 (causal intervention strategy based on dynamic mask, CIDM) 和基于元更新的门控网络结构 (gating network structure based on meta-update,

GNSM). 图 2 为该模型的整体框架.

3.1 基于多头注意力的自适应因果发现方法

针对传统方法难以捕捉多种因果关系且难以随工况变化实时更新的问题,本文提出了一种多头注意力引导的自适应因果发现方法.多头结构使得模型能够并行捕捉变量间可能存在的多重因果关系与非线性依赖,与因果结构学习中多视角检验条件独立性思想一致.同时,注意力权重的实时更新使模型能够自适应地反映工况变化下的因果强度波动,从而实现对工艺过程因果结构的动态发现.

首先,使用多头注意力机制进行动态因果依赖提取.该模型比较节点之间转换后的特征向量以计算注意力分数,注意力分数表示相邻节点的重要性. Softmax 函数对注意力分数进行归一化生成注意力权重.每个节点使用相应的注意力权重聚合来自相邻节点的特征.不同注意力头通过独立的参数映射并行建模变量关系,每个注意力头对应不同的物理或化学依赖模式.例如,一个注意力头聚焦于溶解氧浓度与 pH 值之间的短期依赖关系,另一个则关注物料成分对溶解氧浓度的长期影响.通过这种方式,模型能够捕捉到不同类型的化学和物理依赖关系,从而提升预测精度.然后,对所有注意力头的输出进行平均,获得最终节点表示,输入全连接层.加入自适应更新机制后,注意力权重依赖于节点特征

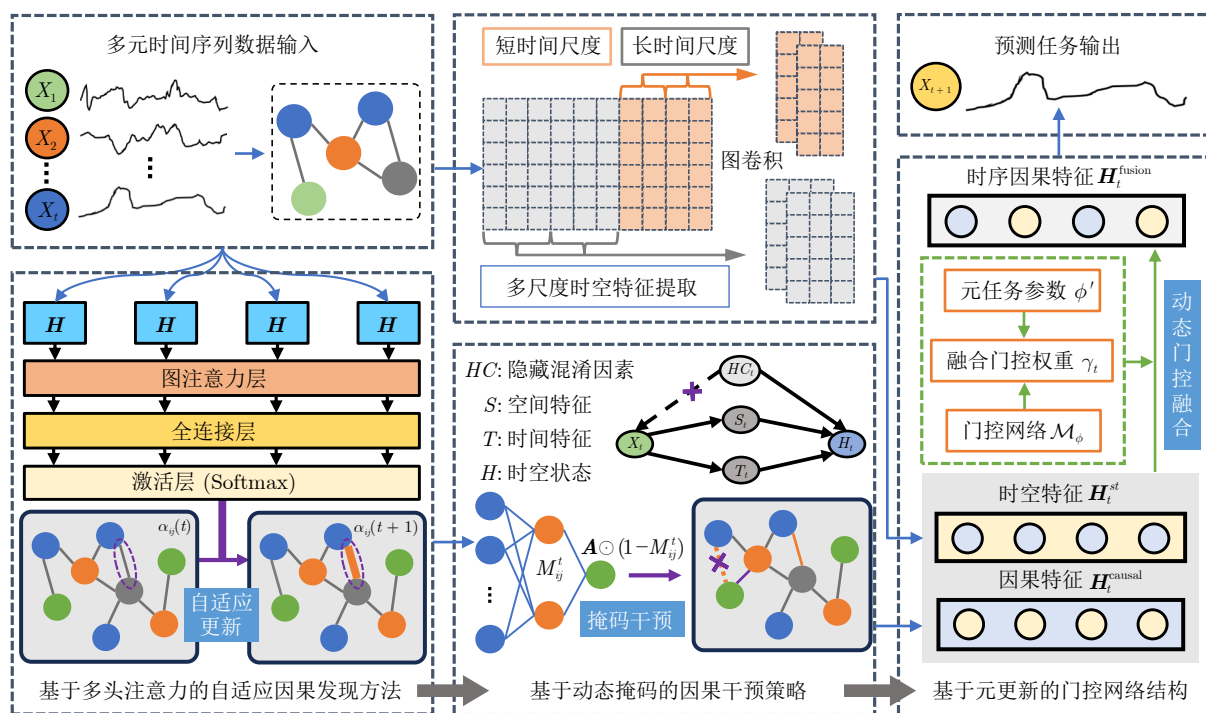


图 2 自适应因果推理预测模型总体框架

Fig. 2 The overall framework of the adaptive causal inference prediction model

和动态信息:

$$\alpha_{ij}(t) = \frac{\exp(M(a^T[W h_i(t) \parallel W h_j(t) \parallel e_i(t)]))}{\sum_{k \in \mathcal{N}(i)} \exp(M(a^T[W h_i(t) \parallel W h_j(t) \parallel e_i(t)]))} \quad (7)$$

其中, $\alpha_{ij}(t)$ 表示节点 v_i 和节点 v_j 之间在 t 时刻的注意力权重; $h_i(t)$ 表示节点 v_i 在 t 时刻的特征向量; $e_i(t)$ 表示节点 v_i 在 t 时刻的环境信息; W 表示学习节点特征转换的权重矩阵; a 表示用于计算注意力的学习参数; $\parallel$ 表示向量的拼接操作; $\mathcal{N}(i)$ 表示节点 v_i 的邻居集合; M 表示 LeakyReLU 激活函数. 多个注意力头计算节点之间的关系, 并基于动态信息进行更新:

$$h'_i(t) = \bigoplus_{k=1}^K \alpha_{ij}^{(k)}(t) W^{(k)} h_j(t) \quad (8)$$

其中, $h'_i(t)$ 表示节点 v_i 在 t 时刻的最终表示; $\alpha_{ij}^{(k)}(t)$ 表示第 k 个注意力头计算的节点 v_i 和节点 v_j 之间的注意力权重; $W^{(k)}$ 表示第 k 个注意力头的权重矩阵; K 表示注意力头的数量; $\bigoplus$ 表示多头注意力的加权求和.

3.2 基于动态掩码的因果干预策略

针对因果关系中虚假相关产生的预测偏差, 设计了一种基于动态掩码的因果干预策略. 该方法通过构建时变的掩码矩阵, 筛选并保留真实因果关系, 从而在特征传播中有效抑制虚假相关的干扰.

首先, 对每一时间步的因果邻接矩阵进行有选择性的掩码干预处理. 掩码构建时同时考虑因果权重大小与因果显著性得分. 因果显著性得分计算基于变量被干预前后模型输出变化幅度的对比, 用以区分不同变量的因果贡献程度. 阈值是因果结构稀疏性约束所确定的理论值, 用于控制因果关系筛选的严格程度. 当二者均超过预设阈值时, 对应因果连接被保留, 保证因果干预结果具有稳健性. 根据因果权重 α_{ij}^t 及因果显著性得分 σ_{ij}^t , 构建掩码矩阵为:

$$M_{ij}^t = \begin{cases} 1, & \alpha_{ij}^t \geq \tau_\alpha \text{ 且 } \sigma_{ij}^t \geq \tau_\sigma \\ 0, & \alpha_{ij}^t < \tau_\alpha \text{ 或 } \sigma_{ij}^t < \tau_\sigma \end{cases} \quad (9)$$

$$\sigma_{ij}^t = \left| y_{\text{do}(x_j)}^t - y^t \right| \quad (10)$$

其中, M_{ij}^t 表示在 t 时刻是否保留从节点 v_i 指向节点 v_j 的因果边; τ_α 为因果强度阈值, 控制因果边的最小权重; σ_{ij}^t 表示该因果关系的因果显著性得分; y^t 为原始条件下模型的预测输出; $y_{\text{do}(x_j)}^t$ 为因果干预后的预测输出; τ_σ 为显著性阈值, 保证被保留的

因果关系可靠. 构建因果干预后的邻接矩阵:

$$\tilde{\mathbf{A}} = \mathbf{A} \odot (1 - M_{ij}^t) \quad (11)$$

其中, $\tilde{\mathbf{A}}$ 为经过干预后的因果邻接矩阵; $\odot$ 表示元素乘法运算. 按掩码筛选因果边的操作等价于在因果推理中执行 do-运算, 即通过外部干预固定某一变量的取值, 并切断其与原有因果父节点之间的依赖关系, 以消除隐藏混淆因素引起的虚假因果路径.

CIDM 能够在每一个时间步自适应调整因果结构, 保留真实的因果关系, 抑制无效信息的传播. 干预后的因果图将作为后续门控图卷积网络的输入, 引导节点间的特征融合过程, 将特征传播过程建立在可解释且稳定的因果依赖上, 有效提升在复杂动态系统下的预测精度.

3.3 基于元更新的门控网络结构

针对时空状态和因果特征融合缺乏自适应调控机制的问题, 建立一种元更新门控网络结构. 该结构通过可学习的门控网络参数自适应生成特征融合权重, 并结合元学习实现门控参数的快速更新和动态优化. 首先, 在每个时间步, 模型分别提取出因果特征 $\mathbf{H}_t^{\text{causal}}$ 和时空特征 $\mathbf{H}_t^{\text{st}}$. 为实现进一步融合, 这两类特征首先在特征维度上拼接得到融合输入为:

$$\mathbf{Z}_t = [\mathbf{H}_t^{\text{causal}} \parallel \mathbf{H}_t^{\text{st}}] \quad (12)$$

其中, $\mathbf{Z}_t$ 为每个节点当前时刻的融合输入表示. 然后使用门控网络 $\mathcal{M}_\phi$ 生成融合门控权重:

$$\gamma_t = \omega(\mathcal{M}_\phi(\mathbf{Z}_t)) \quad (13)$$

其中, ω 为 Sigmoid 函数; γ_t 为融合门控权重, 表示每一维特征的因果占比. 根据门控权重 γ_t , 执行加权融合, 得到最终的融合特征表示为:

$$\mathbf{H}_t^{\text{fusion}} = \gamma_t \odot \mathbf{H}_t^{\text{causal}} + (1 - \gamma_t) \odot \mathbf{H}_t^{\text{st}} \quad (14)$$

其中, $\mathbf{H}_t^{\text{fusion}}$ 表示融合后的节点表示. 为了提升门控网络 $\mathcal{M}_\phi$ 的自适应能力, 引入基于元学习的更新策略. 设当前任务为 $\mathcal{T}$, 对门控网络参数 ϕ 进行一次快速更新, 得到任务自适应参数 ϕ' 为:

$$\phi' = \phi - \rho \nabla_\phi \mathcal{L}_{\mathcal{T}}^{\text{support}}(\phi) \quad (15)$$

其中, $\mathcal{L}_{\mathcal{T}}^{\text{support}}(\phi)$ 表示在支持集上的预测损失函数; ρ 为内层学习率; ϕ' 为当前任务 $\mathcal{T}$ 下的临时优化参数. 然后, 在查询集上使用更新后的参数 ϕ' 生成新的融合门控权重, 并计算其在查询集上的预测损失, 通过该损失反向更新初始参数 ϕ 为:

$$\phi \leftarrow \phi - \beta \nabla_\phi \mathcal{L}_{\mathcal{T}}^{\text{query}}(\phi') \quad (16)$$

其中, β 为外层学习率; $\mathcal{L}_{\mathcal{T}}^{\text{query}}(\phi')$ 表示查询集上的

融合预测误差. 该元更新过程使得门控机制具备快速适应不同任务特征分布的能力, 提升了融合策略的灵活性与泛化性. 最后, 融合特征作为全局表示输入后续预测模块, 用于生成最终任务预测结果.

4 实验结果与讨论

为评估该模型的有效性, 在田纳西州伊士曼 (Tennessee Eastman, TE) 过程数据集^[23]和锌浸出 (Zinc leaching, ZL) 过程数据集上进行了实验验证. 对比较模型、参数分析和实施细节进行了讨论. 同时进行了消融实验来验证 ACIPM 中每个模块的有效性. 模型开发工具为 Torch 1.12.1 和 Python 3.8.0, 操作系统为 Ubuntu 20.04, 硬件配置采用 Intel CORE i7-13700H CPU 和 NVIDIA RTX 4090 GPU, 能够处理复杂的深度学习任务. 具体实验内容将在以下部分介绍.

4.1 实验描述

注意力头的数量设置为 8, 每个头学习输入特征的不同部分. 模型训练从 epoch 0 开始, 总共运行 180 个 epoch, 学习率为 0.001. 目标是对变量之间的各种相关性进行建模, 因此, 在公开的 TE 数据集上进行了相关实验. 同时, ZL 数据集是电子固废处理锌析取过程的关键变量监测. 该数据集中包含了 5 个反应器, 涉及 13 个传感器数据, 包括废酸流量、废料流量、pH 值、氧气流量、溶解氧浓度等. 采样间隔 1 min, 包含 1 000 个时间步长的数据. 数据集被划分为训练集、验证集和测试集, 比例为 6 : 2 : 2. TE 数据集是被广泛用作评估过程控制和监控的基准数据集^[24]. 该数据集代表一个模拟的化学生产过程, 涉及 5 个主要组件, 记录了 8 个可观测变量, 采样间隔 3 min, 包含 480 个时间步长数据. 数据集被划分为训练集、验证集和测试集, 比例为 6 : 2 : 2.

4.2 对比方法和评价指标

通过多步预测任务评估模型的性能, 并与四个现有模型进行比较: DCRNN^[25]、STGCN^[12]、ASTGCN^[26]和 CISTGNN^[22]. DCRNN 通过图卷积捕获空间关系, 并通过结合循环神经网络对时间依赖关系进行建模; STGCN 结合了时间卷积和图卷积, 使其能够捕获时间和空间依赖关系; ASTGCN 增加了自适应机制, 动态调整时空图结构; CISTGNN 结合了因果推理, 允许模型学习时空数据中不同变量之间的因果关系.

通过预测因果关系矩阵与真实因果关系矩阵之间的一致性来评估 ACIPM 学习因果关系的准确

性, 采用了类灰度共生矩阵 (gray level co-occurrence matrices, GLCM) 分析^[27]. 将模型学习得到的因果矩阵视为一种具有空间结构特征的加权邻接表示, 不同元素刻画变量间因果强度的分布模式, 而 GLCM 能够有效刻画二维矩阵中权重分布的整体统计特性, 从而为因果结构的整体质量评估提供量化依据. 为了全面评估多步预测任务的性能, 采用了平均绝对误差 (mean absolute error, MAE)^[28]、平均绝对百分比误差 (mean absolute percentage error, MAPE)^[29]和均方根误差 (root mean square error, RMSE)^[30]. 这些指标可以衡量预测结果与实际观测值的差距, 从不同角度反映预测准确性, 提供了对模型预测性能的详细评估, 指导模型优化和调整.

4.3 实验结果分析

4.3.1 TE 过程数据集

ACIPM 在 TE 数据集上的对比实验结果如表 1 所示. ACIPM 在 MAE、RMSE 和 MAPE 三项指标上均优于其他四种模型. 具体而言, ACIPM 的 MAE 为 8.49, 低于 DCRNN 和 STGCN, 表明 ACIPM 在控制绝对误差和提升预测精度方面表现更佳. 相比之下, ASTGCN 和 CISTGNN 的 MAE 与 ACIPM 相对接近. ACIPM 的 RMSE 低于 DCRNN、STGCN、ASTGCN 和 CISTGNN, 说明该模型在控制大误差方面更具优势, 可以实现更稳定的预测. ACIPM 的 MAPE 为 4.17%, 显著优于 DCRNN 和 STGCN; ASTGCN 和 CISTGNN 的 MAPE 分别为 4.97% 和 5.01%, 均高于 ACIPM, 这表明 ACIPM 在不同尺度数据下能够保持更一致和稳定的预测精度, 在预测任务中展现出显著优势.

表 1 TE 数据集上的对比实验
Table 1 Comparison experiments on the TE dataset

	MAE	RMSE	MAPE (%)
DCRNN	10.75	19.55	6.31
STGCN	10.14	23.42	5.48
ASTGCN	9.15	18.21	4.97
CISTGNN	9.60	19.79	5.01
ACIPM	8.49	17.91	4.17

4.3.2 ZL 过程数据集

ACIPM 在 ZL 数据集上的因果关系热力图对比结果如图 3 所示. 图 3(a) 为真实因果关系, 该因果关系热力图是基于析取过程的工艺机理分析构建

的, 结合了工艺流程、物料传递以及已知的理化作用关系, 用于因果关系建模的对比基准; 图 3(b) 为 CISTGNN 的结果; 图 3(c) 为未因果干预 ACIPM 的结果; 图 3(d) 为完整 ACIPM 的结果. 表 2 为对因果关系热力图进行 GLCM 分析. GLCM 能量用于刻画因果权重分布的集中程度, 能量较高表示因果强度主要集中于少数关键边; GLCM 熵用于衡量分布的不确定性, 熵值较低表明模型对因果关系的判定更一致, 虚假或模糊因果连接较少. 能量与熵从集中性与确定性两个方面描述了因果结构的质量. 其中, 正样本表示真实存在的因果关系, 负样本表示虚假的因果关系. Bhattacharyya 距离衡量正样本分布与负样本分布的可分离度, 值越大, 说明模型区分真实因果边和伪因果边的能力越强.

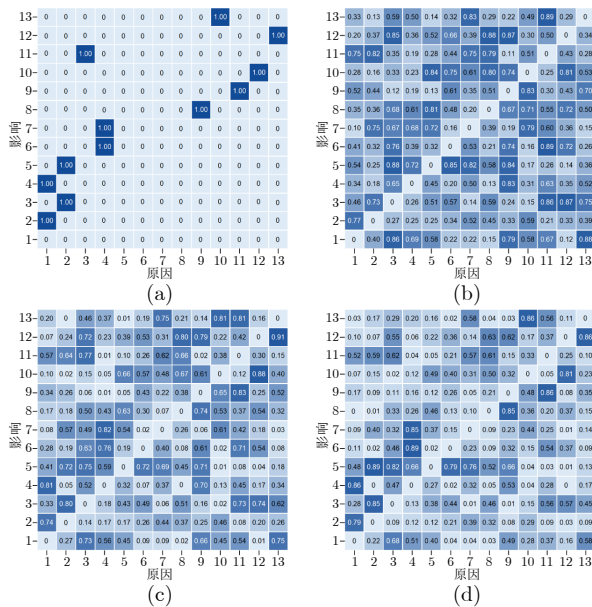


图 3 ZL 数据集的因果关系热力图 ((a) 真实因果关系; (b) CISTGNN; (c) 未因果干预 ACIPM; (d) ACIPM)

Fig.3 Causal relationship heatmaps on the ZL dataset

((a) Truth causal relationship; (b) CISTGNN;
(c) ACIPM without causal intervention;
(d) ACIPM)

基于 GLCM 统计特征可知, 不同模型得到的因果注意力矩阵在正负样本分布特性上存在显著差异: ACIPM 在正样本灰度均值上最高, 而负样本灰度均值最低, 说明该模型能够将注意力集中分配到真实因果关系对应的边上, 并有效抑制非因果边的干扰; CISTGNN 的正负样本灰度均值接近, 表明模型在捕捉因果特征时存在显著混杂, 导致难以区分真实与伪因果关系. 在分布集中性和复杂性方面, ACIPM 的正样本能量最高、熵最低, 而负样本能量相对较低、熵较高, 体现了其在因果关系建模过程

表 2 ZL 数据集上的类 GLCM 特征

Table 2 GLCM-like features on the ZL dataset

	CISTGNN	无 CIDM	ACIPM
正样本灰度均值 ↑	0.50	0.77	0.79
负样本灰度均值 ↓	0.43	0.30	0.21
正样本能量 ↑	0.24	0.42	0.49
负样本能量 ↑	0.14	0.18	0.25
正样本熵 ↓	2.19	1.55	1.42
负样本熵 ↓	2.89	2.62	2.33
Bhattacharyya ↑	0.21	0.86	1.07

注: 无 CIDM 是未因果干预的 ACIPM.

中将正因果关系集中、非因果关系分散, 有助于强化模型对因果依赖的表征能力. 相比之下, 无 CIDM 模型的正负能量与熵分布差异减弱, 显示其在区分因果与非因果特征时存在一定的模糊性. 此外, Bhattacharyya 距离结果进一步佐证了上述结论: ACIPM 达到 1.07, 显著高于无 CIDM 和 CISTGNN, 表明其学习到的正负样本分布几乎不重叠, 可分性最强, 因果特征表达最为清晰. CISTGNN 的距离最低, 仅为 0.21, 显示正负样本分布高度重叠, 模型未能有效捕捉真实的因果结构. 综上所述, ACIPM 在因果注意力分布的均值、集中性、复杂性和分布可分性等多个维度均表现最优, 充分说明引入因果自适应机制能够显著提升模型对真实因果关系的识别能力, 建立准确的动态因果模型.

ZL 数据集中酸碱度和溶解氧浓度的预测曲线如图 4 和图 5 所示, P1、P2 和 D1、D2 均为图 1 所示的监测点. 由图 4 和图 5 可知, 模型在关键工艺参数上整体能够较好地跟踪真实变化趋势. 对于酸碱度, 模型预测能够捕捉主要的峰谷变化, 仅在个别工况发生快速调整的时间段出现幅值欠拟合或相位轻微滞后. 这是由于电子固废原料成分具有显著异质性, 在投料切换或操作参数快速变化时, 反应速率及变量间的因果关系结构可能在短时间内发生变化, 而模型基于历史观测进行预测, 在突变初期存在一定响应滞后. 对于溶解氧浓度, 真实值整体在较窄区间内振荡, 模型预测曲线与真实曲线高度重合, 残差分布较为平稳, 表明所提出方法在刻画稳定因果依赖关系和小幅波动信号方面具有较强能力. 综上所述, 模型在工况突变阶段存在一定局部误差, 但仍能够保持良好的预测性能, 实现了对酸碱度和溶解氧浓度的动态跟踪, 体现出对多种工艺变量因果相关性建模的优势.

ACIPM 在 ZL 数据集上的对比实验结果如表 3 所示. 由表可知 ACIPM 在 MAE、RMSE 和 MAPE 上均表现最优. ACIPM 的 MAE 为 1.34, 显

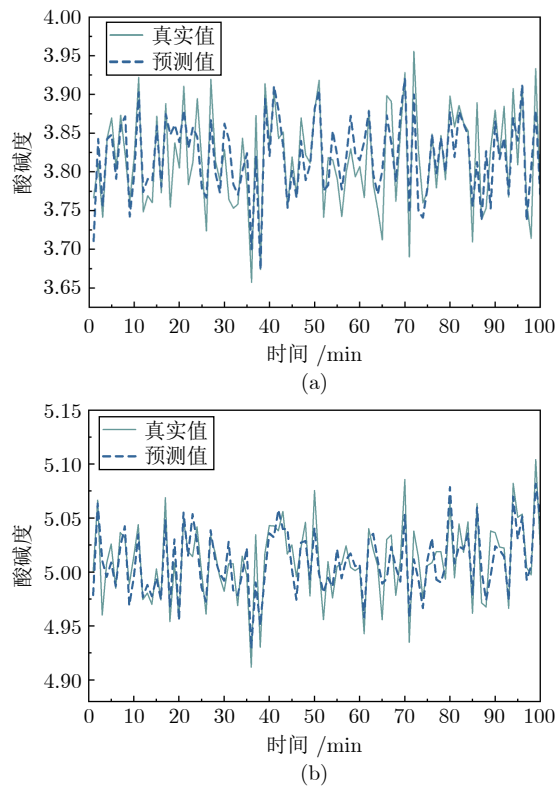


图 4 酸碱度预测曲线 ((a) P1; (b) P2)

Fig.4 Predicted curve of pH value ((a) P1; (b) P2)

著低于其他方法,说明其具有较高的预测精度. 其RMSE 为 2.05, 同样优于对比方法, 进一步表明该方法在面对波动较大的数据样本时依然保持较强的稳定性. 在 MAPE 指标上, ACIPM 达到 5.02%, 显示出对比比例误差的有效控制. 相比之下, 传统的时空图模型 STGCN 以及 ASTGCN 在三个指标上存在较大差距, 表明因果关系学习能有效提高模型预测准确性.

综上所述, ACIPM 方法在因果关系建模与预测任务中均展现出显著优势. 通过自适应机制与因果干预, 模型能够有效抑制虚假相关和混杂因素, 使因果注意力分布更清晰、更具可分性; 同时在预测性能上保持更高的精度与稳定性. 由此可见, 该方法在析取过程建模中不仅提升了因果结构学习的可靠性, 也增强了多变量预测的准确性.

4.4 消融实验研究

消融实验的具体设计方法如表 4 所示. 方案 A 为基础模型, 作为性能对比基准; 方案 B 包括 GNSM, 用于分析特征融合机制对预测性能的影响; 方案 C 包括 ACDM 与 GNSM, 但不包含 CIDM, 以验证在缺乏因果干预条件下, 仅依赖动态因果结构学习的建模效果; 方案 D 同时引入 ACDM 与 CIDM,

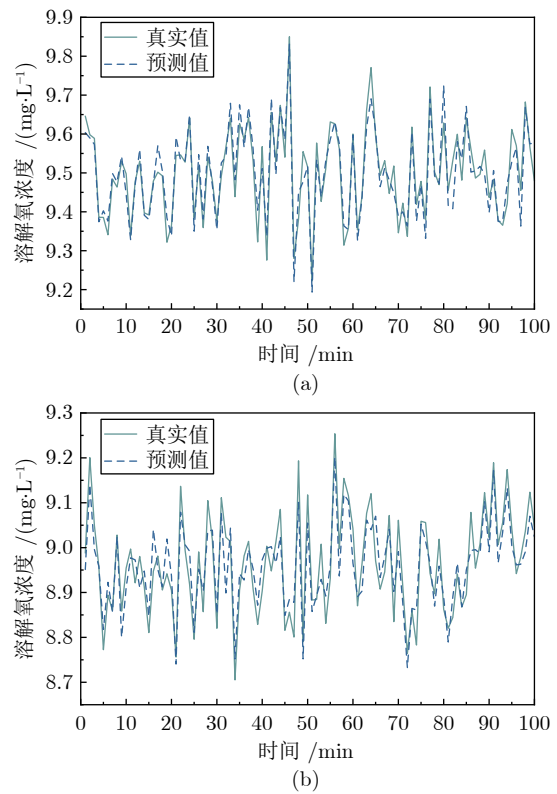


图 5 溶解氧浓度预测曲线 ((a) D1; (b) D2)

Fig.5 Prediction curve of dissolved oxygen concentration ((a) D1; (b) D2)

表 3 ZL 数据集上的对比实验

Table 3 Comparison experiments on the ZL dataset

	MAE	RMSE	MAPE (%)
DCRNN	2.75	4.32	5.41
STGCN	2.84	4.17	5.47
ASTGCN	1.93	3.19	5.81
CISTGNN	1.63	3.40	5.32
ACIPM	1.34	2.05	5.02

表 4 消融实验说明

Table 4 Ablation experiment description

方案	ACDM	CIDM	GNSM
A	×	×	×
B	×	×	✓
C	✓	×	✓
D	✓	✓	×
E	✓	✓	✓

但移除 GNSM, 用于评估在缺乏特征融合机制时因果建模的效果; 方案 E 为完整模型.

ACIPM 在 ZL 数据集与 TE 数据集上的消融实验如图 6 所示, 用于验证各模块的有效性. 从 ZL

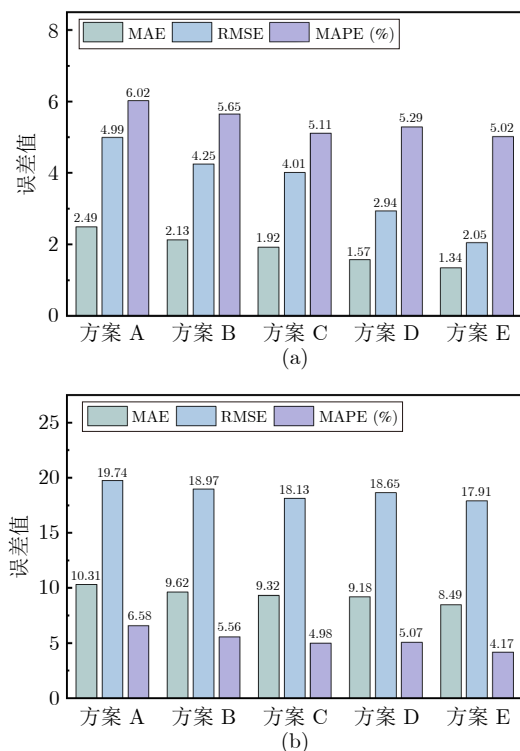


图6 消融实验结果 ((a) ZL数据集; (b) TE数据集)

Fig.6 Ablation experiment results ((a) ZL dataset; (b) TE dataset)

数据集结果来看, 方案 A 的 MAE、RMSE 和 MAPE 分别为 2.49、4.99 和 6.02%, 整体误差较高。在逐步引入模块后, 性能呈现持续提升趋势。相比方案 B, 方案 C 在 MAE 和 RMSE 上明显下降, 表明加入 ACDM 后, 模型能够随工况变化动态刻画变量间依赖结构, 从而提升预测精度。对比方案 C 与方案 E 可知, 加入 CIDM 后性能进一步提升, 说明仅依据相关性学习的因果结构仍包含虚假依赖, 而 CIDM 抑制无效边参与特征传播, 降低了局部误差。对比方案 D 与方案 E 可知, 加入 GNSM 后精度提升, 表明在不同工况下因果特征与时空特征的贡献并不恒定, GNSM 有效解决了这一问题。方案 E 加入全部模块取得最优结果, 在 TE 数据集上也得到一致结论。同时, ZL 数据集中由方案 B 到方案 C 的提升效果大于 TE 数据集, 说明在变量耦合更复杂、结构变化更明显的场景下, 动态因果结构建模精度提升更大。整体结果验证了该模型在析取过程复杂因果关系建模中的有效性。

5 结束语

针对电子固废处理析取过程中由于原料异质性导致的因果关系波动所引起的溶解氧浓度难以准确

预测的问题, 设计了一种自适应因果推理预测模型。该模型通过自适应因果发现、动态掩码因果干预以及元更新门控融合, 动态捕捉并更新工艺变量间的因果结构, 显著提升了反应过程关键参数的预测精度与模型的可解释性。实验结果表明, ACIPM 在复杂析取过程建模中, 能够有效地识别和适应因果关系的动态变化, 表现出比传统方法更高的准确性。将因果推理与时空图神经网络相结合, 为解决现有模型在非平稳数据下因果结构无法动态调整的问题提供了新思路, 促进了因果推理在深度学习中的应用。但模型在极端工况突变及大规模变量场景下的应用仍面临一定挑战, 未来可在更多工业领域数据上进一步验证 ACIPM 的泛化能力。

参考文献

- 1 Foo G, Kara S, Pagnucco M. Artificial learning for part identification in robotic disassembly through automatic rule generation in an ontology. *IEEE Transactions on Automation Science and Engineering*, 2023, **20**(1): 296–309
- 2 Han H G, Liu Y M, Li F Y, Du Y P. Sparse convolutional model with semantic expression for waste electrical appliances recognition. *Science China Technological Sciences*, 2024, **67**(9): 2881–2893
- 3 Han H G, Zhang Q Y, Li F Y, Du Y P, Gu Y F, Wu Y F. Metallic product recognition with dual attention and multi-branch residual blocks-based convolutional neural networks. *Circular Economy*, 2022, **1**: Article No. 100014
- 4 Cao Li-Feng, Yan Gao-Wei, Xiao Shu-Yi, Dong Zhen-Zhu, Dong Ping. Time series prediction method based on domain adaptation physics-informed neural network. *Acta Automatica Sinica*, 2025, **51**(6): 1329–1346
(曹力丰, 阎高伟, 肖舒怡, 董珍柱, 董平. 基于域适应物理信息神经网络的时间序列预测方法. *自动化学报*, 2025, **51**(6): 1329–1346)
- 5 Zhu S Y, Zheng J W, Ma Q L. MR-Transformer: Multiresolution Transformer for multivariate time series prediction. *IEEE Transactions on Neural Networks and Learning Systems*, 2025, **36**(1): 1171–1183
- 6 Duan C M, Wu Z C, Zhu L, Xu X R, Zhu J M, Wei Z Q, et al. DSTN: Dynamic spatio-temporal network for early fault warning in chemical processes. *Knowledge-Based Systems*, 2024, **296**: Article No. 111892
- 7 Han H G, Sun C X, Wu X L, Yang H Y, Zhao D Z. Self-organizing stacked type-2 fuzzy neural network with rule generation. *IEEE Transactions on Neural Networks and Learning Systems*, 2025, **36**(6): 10128–10142
- 8 He Z H, Xu L Y, Yu J, Wu X R. Dynamic multi-fusion spatio-temporal graph neural network for multivariate time series forecasting. *Expert Systems With Applications*, 2024, **241**: Article No. 122729
- 9 Fang H, Wang H S, Gao Y, Zhang Y G, Bu J J, Han B, et al. InsGNN: Interpretable spatio-temporal graph neural networks via information bottleneck. *Information Fusion*, 2025, **119**: Article No. 102997
- 10 Qi X Y, Hu W J, Li B S, Han K. STGNN-FAM: A traffic flow prediction model for spatiotemporal graph networks based on fusion of attention mechanisms. *Journal of Advanced Transportation*, 2023, **2023**(1): Article No. 8880530
- 11 Xu Y, Zhang C, Luo Y, Ke W, Zhu Q X, He Y L. Multiattention spatiotemporal fusion graph neural network for chemical process fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, 2025, **74**: Article No. 3530113
- 12 Yu B, Yin H T, Zhu Z X. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 2023, **2023**(1): Article No. 8880530

- cial Intelligence. Stockholm, Sweden: IJCAI, 2018. 3634–3640
- 13 Qian L P, Zuo Q, Liu H G, Zhu H. Multivariate time series classification based on spatial-temporal attention-dynamic graph neural network. *Applied Intelligence*, 2025, **55**(2): 1–15
 - 14 Qu Y H, Jia X L, Guo J H, Zhu H R, Wu W B. MSSTGNN: Multi-scaled spatio-temporal graph neural networks for short- and long-term traffic prediction. *Knowledge-Based Systems*, 2024, **306**: Article No. 112716
 - 15 Sun H Z, Tang X H, Lu J G, Liu F J. Spatio-temporal graph neural network for traffic prediction based on adaptive neighborhood selection. *Transportation Research Record*, 2024, **2678**(6): 641–655
 - 16 Zheng Y, Koh H Y, Jin M, Chi L H, Phan K T, Pan S R. Correlation-aware spatial-temporal graph learning for multivariate time-series anomaly detection. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, **35**(9): 11802–11816
 - 17 Li Z L, Yu J, Zhang G W, Xu L Y. Dynamic spatio-temporal graph network with adaptive propagation mechanism for multivariate time series forecasting. *Expert Systems With Applications*, 2023, **216**: Article No. 119374
 - 18 Nauta M, Bucur D, Seifert C. Causal discovery with attention based convolutional neural networks. *Machine Learning and Knowledge Extraction*, 2019, **1**(1): 312–340
 - 19 Zhu Q H, Chen S W, Guo T, Lv Y S, Du W B. A spatio-temporal approach with self-corrective causal inference for flight delay prediction. *IEEE Transactions on Intelligent Transportation Systems*, 2024, **25**(12): 20820–20831
 - 20 Miraki A, Parvinian P, Arghandeh R. Electricity demand forecasting at distribution and household levels using explainable causal graph neural network. *Energy and AI*, 2024, **16**: Article No. 100368
 - 21 Wang H, Liu R N, Ding S X, Hu Q H, Li Z X, Zhou H K. Causal-trivial attention graph neural network for fault diagnosis of complex industrial processes. *IEEE Transactions on Industrial Informatics*, 2024, **20**(2): 1987–1996
 - 22 Luo X X, Yin W, Li Z. Spatio-temporal graph neural network with hidden confounders for causal forecast. In: Proceedings of the 32nd International Conference on Case-Based Reasoning. Yucatán, Mexico: Springer, 2024. 157–169
 - 23 Song P Y, Zhao C H, Huang B. Addressing heterogeneous time-frequency causality: Source consistency exploring for industrial root cause alignment and diagnosis. *IEEE Transactions on Cybernetics*, 2025, **55**(3): 1107–1120
 - 24 Song P Y, Zhao C H, Huang B. MPGE and rootrank: A sufficient root-cause characterization and quantification framework for industrial process faults. *Neural Networks*, 2023, **161**: 397–417
 - 25 Li Y G, Yu R, Shahabi C, Liu Y. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In: Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada: ICLR, 2018. 52–57
 - 26 Guo S N, Lin Y F, Feng N, Song C, Wan H Y. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence. Hawaii, USA: AAAI, 2019. 922–929
 - 27 Portes L, Pirot G, Nzikou M M, Giraud J, Lindsay M, Jessell M, et al. Feature-fusion enhanced t-SNE image atlas for geophysical features discovery. *Scientific Reports*, 2025, **15**(1): Article No. 17152
 - 28 Li J C, Fan Y, Wang Z Y, Wang Y J. A graph neural network with dual-stage feature aggregation for industrial soft sensors. *IEEE Transactions on Instrumentation and Measurement*, 2025, **7**: Article No. 3557410
 - 29 Zheng J L, Saxena D, Cao J N. COIN-GNN: Inductive spatial-temporal prediction for continuous distribution shifts via graph neural networks. *IEEE Transactions on Knowledge and Data Engineering*, 2025, **37**(11): 6320–6333
 - 30 Wang Y L, Wang X L, Zhou J Y, Yang C H, Yang Y. Long se-

quence multivariate time-series forecasting for industrial processes using SASGNN. *IEEE Transactions on Industrial Informatics*, 2024, **20**(10): 12407–12417



韩红桂 北京工业大学教授. 主要研究方向为典型资源循环利用过程智能优化控制, 神经网络结构设计 with 优化. 本文通信作者.

E-mail: rechardhan@bjut.edu.cn

(**HAN Hong-Gui** Professor at Beijing University of Technology.)

His research interests include intelligent optimization and control of typical resource recycling processes, structure design and optimization of neural networks. Corresponding author of this paper.)



任晨昊 北京工业大学环境科学与工程学院博士研究生. 主要研究方向为电子固废析取过程智能建模, 神经网络设计与优化.

E-mail: renchenhao@emails.bjut.edu.cn

(**REN Chen-Hao** Ph.D. candidate at the College of Environmental

Science and Engineering, Beijing University of Technology. His research interests include intelligent modeling of electronic waste disjunction process, and neural network design and optimization.)



李方昱 北京工业大学信息科学技术学院教授. 主要研究方向为智能感知与建模, 复杂系统分析与控制, 分布式智能检测优化.

E-mail: fangyu.li@bjut.edu.cn

(**LI Fang-Yu** Professor at the School of Information Science and

Technology, Beijing University of Technology. His research interests include intelligent perception and modeling, complex system analysis and control, and distributed intelligent detection and optimization.)



刘 峥 北京工业大学信息科学技术学院讲师. 主要研究方向为复杂系统智能特征建模, 智能优化控制.

E-mail: liuzheng@bjut.edu.cn

(**LIU Zheng** Lecturer at the School of Information Science and Techno-

logy, Beijing University of Technology. His research interests include intelligent feature modeling of complex systems and intelligent optimization control.)