

# 深度学习在工业废水预测与控制策略中的研究

李坤 李京京 徐军 田凤蓉 王开春

中蓝连海设计研究院有限公司

DOI:10.12238/eep.v7i8.2199

**[摘要]** 工业废水污染物的预测和控制对环境保护至关重要。本文研究了基于深度学习的工业废水污染物预测技术及优化对策,探讨了其在高效性、精确性、自适应性、可扩展性和实时智能化方面的应用特点。分析了当前技术面临的数据质量、获取挑战、模型复杂度、计算资源消耗及泛化能力不足等问题。提出了通过数据增强、多源数据融合、模型优化和专家系统结合多模型融合的优化策略。研究表明,深度学习技术在该领域有重要应用前景,但仍需优化提升性能。

**[关键词]** 工业废水; 深度学习; 污染物预测; 数据增强

中图分类号: X171.5 文献标识码: A

## Research on prediction and control strategy of industrial wastewater pollutants based on deep learning

Kun Li Jingjing Li Jun Xu Fengrong Tian Kaichun Wang

Bluestar Lehigh Engineering institute Co., Ltd

**[Abstract]** The prediction and control of industrial wastewater pollutants is very important for environmental protection. In this paper, the prediction technology of industrial wastewater pollutants based on deep learning and its optimization strategy are studied, and its application characteristics in the aspects of high efficiency, accuracy, adaptability, scalability and real-time intelligence are discussed. The problems of data quality, acquisition challenge, model complexity, computing resource consumption and insufficient generalization ability are analyzed. An optimization strategy combining multi-model fusion with data enhancement, multi-source data fusion, model optimization and expert system is proposed. Research shows that deep learning technology has important application prospects in this field, but it still needs to be optimized to improve its performance.

**[Key words]** industrial wastewater; Deep learning; Pollutant prediction; Data enhancement

### 引言

随着工业化进程加快,工业废水排放量不断增加,成为环境污染的主要来源之一。废水中的有害污染物不仅威胁水生态系统,还可能危及人类健康。传统的物理和化学检测方法虽然能提供准确数据,但过程繁琐且缺乏实时性,难以满足现代废水处理需求。深度学习技术凭借其强大的数据处理能力和非线性建模优势,在污染物预测中表现优异,不仅提高了预测精度,还实现了实时监测与智能控制,为环境保护提供新方案。本文将探讨深度学习在工业废水污染物预测中的应用特点、存在问题及优化对策,为环境污染控制提供参考。

### 1 基于深度学习的工业废水污染物预测特点

#### 1.1 数据驱动的高效性与精确性

深度学习模型的应用能够大幅提升工业废水污染物预测的效率和准确性。传统的污染物预测方法往往依赖于线性回归或统计分析,其对数据的要求高且在处理非线性关系时存在明显

的局限性。深度学习通过构建多层神经网络,可以在不依赖明确模型假设的情况下,从海量数据中自动提取特征并学习复杂的非线性关系。这种数据驱动的方式使得模型能够处理不同种类的污染物和废水类型,对复杂的污染物浓度变化具有高度的敏感性<sup>[1]</sup>。高效的数据处理能力不仅提高了预测的速度,还减少了人为干预的必要性,从而降低了操作复杂性和潜在的误差。深度学习模型可以通过不断的训练和更新来优化预测结果,实现对污染物的精准监控,有效降低工业废水中有害物质的排放,缓解环境污染问题。

#### 1.2 自适应性及可扩展性

基于深度学习的工业废水污染物预测系统具有很强的自适应性和可扩展性,这使得它能够应对各种不同的工业废水处理场景。深度学习模型通过不断的训练能够自适应不同类型的废水,识别其中特定的污染物种类及其浓度,这种自适应性使得模型在应用过程中具备较高的灵活性。深度学习方法不局限于某

一种特定的污染物,可以通过调整输入的数据和模型结构来处理不同的污染物组合,这为模型在不同行业中的广泛应用提供了可能<sup>[2]</sup>。深度学习模型的可扩展性也体现在其对数据量的处理能力上。随着工业生产的规模化和数据采集技术的进步,废水处理系统需要处理的数据量越来越大。深度学习模型可以轻松扩展,处理海量数据,同时保持预测的精度和效率,为大规模环境监控提供了技术支持,从而在更大范围内防控工业废水污染,保护生态环境。

### 1.3 预测结果的实时性与智能化

实时性和智能化是深度学习在工业废水污染物预测中的两大显著优势。这种预测模型能够快速处理和分析大量实时数据,提供即刻的污染物浓度预测,为工业废水处理提供及时的参考。在现代工业生产中,废水中污染物的浓度随时间不断变化,如果不能及时发现和控制,会对环境造成严重的污染。深度学习模型通过实时数据分析和预测,可以在污染物超标之前就发出预警,指导废水处理系统进行必要的调整<sup>[3]</sup>。深度学习模型与物联网技术相结合,可以实现废水处理过程的全自动化和智能化。通过将深度学习算法嵌入传感器和控制系统中,整个废水处理流程可以在无人干预的情况下自主运行,实时优化处理策略,减少人力成本和误操作风险。智能化的废水处理不仅提高了系统的响应速度和处理效率,还显著降低了对环境的污染,有助于实现工业废水的绿色管理和可持续发展。

## 2 基于深度学习的工业废水污染物预测存在问题

### 2.1 数据质量与数据获取挑战

在基于深度学习的工业废水污染物预测中,数据质量与数据获取是两大关键挑战。深度学习模型的性能高度依赖于训练数据的质量和数量,而在工业废水处理中,获取高质量的数据常常面临诸多困难。不同工业过程产生的废水成分复杂多变,采集的数据存在噪声、不完整或异常值,这会直接影响模型的训练效果和预测精度。获取足够的历史数据进行模型训练也是一个难题<sup>[4]</sup>。数据采集设备和技术的限制、数据收集成本的高昂,以及隐私和安全问题,都导致数据来源有限。数据的不平衡问题在工业废水处理中也较为常见,某些污染物在废水中出现的频率较低,导致模型在检测这些污染物时表现不佳。数据质量的不足不仅限制了深度学习模型的有效性,还导致错误的污染物预测和环境管理决策,进一步加剧环境污染问题。

### 2.2 模型的复杂度与计算资源消耗

深度学习模型的复杂度和计算资源消耗是工业废水污染物预测中不可忽视的问题。深度学习模型通常由多层神经网络组成,模型的复杂性使得它在处理海量数据时能够表现出强大的学习能力。这种复杂性也带来了高昂的计算成本。模型的训练和推理时间较长,尤其在处理实时数据时,难以满足工业废水管理的实时性需求。高计算资源消耗不仅增加了企业的运营成本,还限制了深度学习模型在资源受限环境中的应用<sup>[5]</sup>。复杂的模型结构也增加了调试和优化的难度,需要专业知识和技术支持,这在一定程度上阻碍了深度学习技术在工业废水处理中的广泛应用。

### 2.3 预测模型的泛化能力与适应性

深度学习模型的泛化能力和适应性在工业废水污染物预测中确实面临着重大挑战。泛化能力指的是模型在处理未见过的数据时,依然能够保持高精度预测的能力。然而,由于工业废水的成分和污染物种类会随着工艺流程的变化而发生显著变化,模型在训练过程中通常只能接触到一部分污染物类型或组合。这种局限性导致模型在实际应用中,特别是面对从未见过的新污染物时,可能无法准确预测,进而影响其实际应用效果。模型的适应性问题也不容忽视。工业废水的成分在不同工业部门、不同地区乃至不同时间段内都有可能存在巨大的差异。单一模型在这些情况下难以同时适应所有的场景和复杂情况,导致模型在不同应用环境中的表现不一致。由于这些差异性,深度学习模型可能在某些特定场景下表现良好,而在其他场景下则难以达到预期效果。

## 3 基于深度学习的工业废水污染物预测优化对策

### 3.1 数据增强与多源数据融合

为了解决深度学习模型在工业废水污染物预测中所面临的数据质量和数据获取问题,采用数据增强和多源数据融合是行之有效的优化策略。数据增强技术通过对现有数据进行各种操作,如随机扰动、旋转、翻转、缩放和噪声添加等,生成大量的新数据样本。这些操作不仅可以扩展数据集的规模,还能显著增加数据的多样性,从而改善模型的泛化能力,增强其在处理未见过的样本时的表现。此外,数据增强还可以平衡数据集中的类别分布,减少偏差,使模型更具鲁棒性。

多源数据融合则是通过整合来自不同来源或不同时间段的数据,将其合并为一个更加全面且精确的训练数据集。这种策略能够充分利用各个数据源的优点,弥补单一数据来源的不足,从而提高数据的覆盖范围和可信度。尤其在工业废水处理过程中,废水成分复杂且多变,融合多源数据能够捕捉到更多潜在的污染物变化趋势,使得模型能够更好地应对实际场景中的复杂性。这些数据处理策略的综合实施,能够显著提高深度学习模型在工业废水污染物预测中的准确性和鲁棒性,确保模型在不同的环境和条件下均能保持良好的预测效果。这不仅有助于提升污染物检测的精度,也为更有效地防控环境污染提供了技术支持,推动了工业废水管理的智能化与可持续发展。

### 3.2 模型优化与轻量化策略

针对深度学习模型复杂度高和计算资源消耗大的问题,可以采用模型优化与轻量化策略来提升工业废水污染物预测的效率。模型优化可以通过调整和改进模型结构,减少参数数量和网络层数,以降低计算复杂性。这种优化可以通过选择合适的激活函数、正则化方法以及采用更有效的训练算法来实现,从而提高模型的计算效率,减少训练和推理所需的时间。模型优化还可以包括特征选择和降维技术,通过减少冗余信息来简化输入数据,从而进一步提升计算性能。

轻量化策略是解决深度学习模型在资源受限环境中运行的有效手段。通过网络剪枝,可以去除神经网络中不必要的连接和

节点,从而减少模型的体积和计算量。参数量化将模型参数从高精度的浮点数转化为低精度的整数表示,这在保持模型性能的同时显著减少了存储和计算的资源需求。知识蒸馏则是通过训练一个小型模型来模仿大型模型的行为,从而在不显著降低性能的情况下实现模型的轻量化。迁移学习也是一种有效的策略。通过将预先在大规模数据集上训练好的模型迁移到工业废水数据上进行微调,可以在保持模型性能的同时减少训练时间和数据需求。这种方法特别适用于工业废水处理领域,因为它能够快速适应特定场景的要求,减少了从头开始训练模型的成本和时间。

### 3.3 结合专家系统与多模型融合

结合专家系统与多模型融合是提升深度学习模型泛化能力和适应性的有效方法,特别是在复杂的工业废水污染物预测领域。在此背景下,专家系统的引入具有重要意义。通过编码领域专家的知识 and 经验规则,专家系统可以在深度学习模型的预测结果基础上提供额外的校正和解释功能。这不仅可以增强模型的解释能力,还能为预测结果提供有力的理论支撑,帮助识别模型可能未能充分捕捉的细微特征或异常情况。专家系统的知识库随着时间的推移和数据的积累可以不断更新,使其具备动态适应新的工业废水污染物的能力。

多模型融合则通过整合多个不同模型的预测结果,进一步提高了预测的稳定性和准确性。这种方法利用了不同模型在处理特定数据时的优势,通过加权平均或投票决策等集成学习方法,将各个模型的输出结果进行综合考虑。这样做的好处是可以显著降低单一模型可能出现的偏差或误差,增强系统对不同污染物特征的识别能力。尤其是在处理复杂的工业废水样本时,多模型融合能够提供更加稳健的预测,避免因某一模型对特定数据的适应不佳而导致的预测失误。这种多层次的控制策略不仅有助于提高模型对各种工业废水场景的适应性,还能在一定程度上弥补深度学习模型在面对新污染物时存在的泛化能力不足的问题。通过结合专家系统与多模型融合,能够实现更全面、

精确的污染物预测,从而为工业废水的管理和环境保护提供更加全面和可靠的技术支持。

## 4 总结

本文研究了基于深度学习的工业废水污染物预测技术及其优化对策,旨在提升预测的精度和效率,为环境污染控制提供有效手段。研究表明,深度学习技术在处理工业废水中的复杂污染物组合时展现出显著优势,尤其是在实时性和智能化方面。数据质量和获取难题、模型复杂度和计算资源消耗问题,以及模型泛化能力的不足,都是目前技术应用中面临的主要挑战。本文提出了多种优化对策,包括数据增强与多源数据融合、模型优化与轻量化策略,以及结合专家系统与多模型融合。通过这些措施,可以有效提升深度学习模型的性能,使其更好地适应工业废水污染物的多样性和复杂性,为实现绿色和可持续的工业生产提供技术支持。这一研究为未来深入探索和推广基于深度学习的环境污染物预测和控制策略奠定了基础。

## [参考文献]

- [1]张宇.基于组合模型的旅游交通流预测方法研究与应用[D].兰州大学,2022.
- [2]廖海斌,袁理,龚颖巍.基于多尺度时空图神经网络的污染物浓度预测[J].控制与决策,2024(004):039.
- [3]郭香兰,王立,金学波,等.基于生成对抗网络和深度森林结合的粮食加工过程污染物小样本数据扩充及预测[J].食品科学,2024,45(12):22-30.
- [4]张鑫磊,张冬峰.基于多通道长短期记忆网络的PM<sub>2.5</sub>小时浓度预报[J].环境科学研究,2022,35(12):2685-2692.
- [5]马吉伟.基于强化深度学习的城市环境空气污染监测与预警方法研究[J].环境科学与管理,2023,48(10):116-120.

## 作者简介:

李坤(1988—),男,汉族,安徽省滁州市人,硕士研究生,高级工程师,研究方向:工业废水生物强化、资源化及综合治理。