

人工智能驱动下的大坝安全监测数据异常识别算法研究

刘志萍

云南航安工程检测有限公司 云南 昆明 661100

摘要: 本文系统探讨了人工智能驱动下大坝安全监测数据异常识别的核心算法体系,深入剖析了基于统计学习、机器学习与深度学习的三类主流异常检测方法的理论基础、适用场景与内在局限;重点论述了面向大坝监测数据特性的算法适配性设计原则,包括时空耦合建模、多源信息融合、小样本学习与可解释性增强等关键技术路径;并在此基础上,前瞻性地分析了AI驱动异常识别算法在工程化落地过程中所面临的数据质量、模型泛化、实时性保障与伦理合规等现实挑战。研究表明,构建融合物理机理与数据驱动的混合智能模型,是提升大坝安全监测异常识别精度、鲁棒性与可信度的未来发展方向。本文旨在为相关领域的理论研究与工程实践提供系统性参考。

关键词: 人工智能; 大坝安全监测; 异常识别; 机器学习; 深度学习; 数据融合; 可解释性

引言

大坝作为调控水资源、防洪减灾等的关键工程,其结构安全关乎国家水安全战略。历史上重大溃坝事故警示我们,持续精准高效监测预警大坝结构状态是防灾首要防线。传统监测数据分析依赖人工经验、阈值报警及简单统计方法,处理低维、规则、平稳数据有一定效果,但面对现代大坝监测系统产生的复杂数据流,局限性凸显:高频数据使人工筛查效率低易出错,多环境荷载耦合致大坝行为模式非线性、非平稳,传统静态阈值法误报漏报多。而人工智能中机器学习与深度学习凭借强大能力,为从复杂数据中挖掘异常信号开辟新途径。AI驱动异常识别能动态构建模型,捕捉早期异常、感知突发险情。本文旨在超越具体场景,从算法理论层面梳理探讨适用于大坝安全监测数据异常识别的人工智能方法体系。

1 大坝安全监测数据异常的内涵与特征

1.1 数据异常的内涵

在探讨识别算法之前,必须首先明确“异常”在大坝安全监测语境下的具体含义。此处的异常并非指数值上的简单离群点,而是指监测数据序列中偏离其在特定环境条件下预期正常行为模式的观测值或模式,这种偏离可能预示着结构性能的退化、潜在损伤的发生或外部极端事件的冲击。

1.2 数据异常的类型

根据成因与表现形式,大坝监测数据异常大致可分为三类:(1)结构性异常:由大坝本体或其基础的物理损伤(如裂缝扩展、渗漏通道形成、材料老化、基础沉降等)引起。这类异常通常表现为监测量(如位移、渗压、应变)随时间呈现非预期的趋势性变化、突变或周期性模式的畸变,是安全预警最关注的核心对象。(2)

环境诱发异常:由超出常规范围的极端环境荷载(如特大洪水、强震、极端温差)引起。虽然大坝结构本身可能未受损,但其响应超出了历史正常波动的包络线。区分此类异常与结构性异常至关重要,前者属于可恢复的弹性响应,后者则意味着不可逆的损伤^[1]。(3)系统性异常:源于监测系统自身的故障,如传感器漂移、失效、通信中断或数据采集错误。这类异常虽不直接反映结构状态,但会严重污染数据,干扰对真实结构异常的判断,因此也需被有效识别和剔除。

1.3 特性与挑战

大坝监测数据的固有特性给异常识别带来了巨大挑战:(1)强时空耦合性:大坝是一个连续的三维实体,任一测点的响应都与其空间邻近测点的状态紧密相关,同时又随时间动态演化。忽略这种时空关联性,孤立地分析单个时间序列,极易导致误判。(2)多源异构性:监测数据涵盖几何量(位移、挠度)、物理量(渗压、水温、气温)、力学量(应力、应变)等多种类型,采样频率、精度、单位各不相同,且彼此间存在复杂的物理因果关系。(3)非平稳性与时变性:大坝的“正常”行为模式并非一成不变,它会随水库调度(水位变动)、季节更替(温度循环)及材料徐变(时效效应)而缓慢演化。一个有效的异常识别模型必须能自适应地跟踪这种“正常”的漂移。(4)小样本与不平衡性:真实的结构性异常事件在长期监测记录中极为罕见,导致可用于训练的正样本(异常数据)数量极少,形成了典型的类别极度不平衡问题。这些特性决定了通用的异常检测算法不能直接套用,必须进行深度的领域适配与算法创新。

2 人工智能驱动的异常识别核心算法体系

针对上述挑战,AI社区发展了一系列异常识别算法,

可归纳为三大范式。

2.1 基于统计学习的异常识别方法

此类方法假设正常数据服从某种已知或可推断的概率分布，通过建立该分布模型来识别不符合该分布的观测点。(1) 多元高斯模型：对于多个相互关联的监测变量，可构建联合高斯分布。通过计算新观测点的马氏距离，可以衡量其偏离整体数据中心的程度。该方法能有效捕捉变量间的线性相关性，但对于大坝响应中的强非线性关系则显得力不从心。(2) 高斯过程回归 (GPR)：GPR是一种非参数贝叶斯方法，特别适合于对具有不确定性和噪声的时间序列进行建模。它可以为每个预测点提供一个置信区间，落在置信区间之外的观测即可视为异常。GPR的优势在于其天然的不确定性量化能力，非常适合用于描述大坝在环境荷载作用下的正常响应带。(3) 隐马尔可夫模型 (HMMs)：HMM将大坝的“健康状态”视为一个不可观测的隐状态序列，而监测数据则是该隐状态下的观测输出。通过学习状态转移概率和观测发射概率，HMM可以推断当前最可能的隐状态^[2]。当系统长时间处于一个代表“异常”的隐状态时，即可发出警报。这种方法擅长处理具有状态切换特性的序列数据。统计学习方法通常具有良好的数学可解释性，但其性能高度依赖于对数据分布先验假设的准确性，在面对大坝复杂非线性动力学系统时，其表达能力受限。

2.2 基于机器学习的异常识别方法

机器学习方法通过从数据中直接学习决策边界或重构规则，摆脱了对特定概率分布的强假设。(1) 一类支持向量机 (OC-SVM)：OC-SVM是专门针对单类分类（即只有正常样本可用）问题设计的。它试图在高维特征空间中找到一个能包含所有（或绝大多数）正常样本的最小超球体，位于球体外的点即为异常。通过核技巧，OC-SVM能有效处理非线性问题，是小样本异常检测的经典利器。(2) 孤立森林 (iForest)：iForest基于一个直观思想：异常点因其稀有性和与正常点的差异性，能够在随机分割的过程中被更快地“孤立”出来。算法通过构建多棵随机隔离树，并计算样本的平均路径长度来度量其异常程度。iForest计算高效，对高维数据和大数据集有很好的扩展性，且无需对数据分布做任何假设。(3) 局部异常因子 (LOF)：LOF通过比较一个点与其邻居的局部密度来判断其异常程度。如果一个点的局部密度远低于其邻居，则认为它是异常点。这种方法对识别局部簇中的异常点非常有效，可以捕捉到全局统计方法可能忽略的细微异常。这些无监督或半监督的机器学习算法，以其强大的非线性建模能力和对数据分布的弱依赖性，

成为大坝异常识别的重要工具。

2.3 基于深度学习的异常识别方法

深度学习利用深层神经网络强大的表示学习能力，能够自动从原始数据中提取多层次、抽象化的特征，从而实现端到端的异常检测。(1) 自编码器 (AE) 及其变体：AE由编码器和解码器组成，旨在学习一个恒等映射。在仅使用正常数据训练后，AE能够很好地重构正常输入，而对异常输入的重构误差会显著增大。通过设定重构误差阈值即可识别异常。变体如变分自编码器 (VAE) 引入了概率生成机制，去噪自编码器 (DAE) 则增强了模型的鲁棒性。(2) 长短期记忆网络 (LSTM)：LSTM是专为处理长序列依赖问题设计的循环神经网络 (RNN)。它可以学习大坝监测序列中复杂的长期动态模式。通过训练一个序列到序列 (Seq2Seq) 的LSTM模型来预测未来时刻的正常值，实际观测值与预测值之间的残差即可作为异常指标^[3]。(3) 图神经网络 (GNNs)：GNN是处理图结构数据的强大工具。将大坝的传感器网络抽象为图（节点为传感器，边为物理连接或统计相关性），GNN能够同时聚合空间邻域信息和时间动态信息，实现真正的时空联合建模。这对于捕捉由局部损伤引发的、沿结构传播的异常信号模式具有天然优势。深度学习方法代表了当前异常检测的前沿，其卓越的性能是以牺牲一定的模型可解释性和增加计算复杂度为代价的。

3 面向大坝监测特性的算法适配与优化

3.1 时空耦合建模

单一的时间序列模型或空间快照模型均无法完整刻画大坝行为。理想的模型应能同步处理时空维度。除了前述的GNN，还可以采用时空卷积网络 (ST-ConvNet)，利用三维卷积核同时在时间和空间上进行特征提取；或者构建时空注意力机制，让模型能够自适应地关注对当前预测或异常判断最重要的时空区域。

3.2 多源信息融合

有效的异常识别必须整合来自不同物理量的信息。一种策略是构建多模态深度学习架构，例如，为每种数据类型（位移、渗压、温度）设计独立的特征提取子网络（如CNN或LSTM），然后将它们的高层特征进行融合（拼接、加权求和或通过注意力机制），再送入最终的异常判别模块。另一种策略是利用物理知识引导融合，例如，将水位、温度等环境变量作为外部输入引入到预测模型（如LSTM）中，使模型学习的是“在给定的环境条件下”的结构响应，从而更准确地分离出由结构自身变化引起的异常。

3.3 小样本学习与主动学习

针对异常样本稀缺的问题,可以采用迁移学习,将在其他类似大坝或模拟数据上预训练的模型,迁移到目标大坝进行微调^[4]。此外,主动学习框架可以被引入:模型在运行过程中,对那些预测置信度低的“疑似”异常样本,主动请求专家进行标注,从而以最少的人工干预成本,持续优化模型性能。

3.4 可解释性与可信AI

在安全攸关的工程领域,模型的“黑箱”特性是其应用的重大障碍。必须增强模型的可解释性(XAI)。例如,可以利用梯度加权类激活映射(Grad-CAM)等技术,可视化深度模型在做出异常判断时所依据的原始数据区域;或者采用基于注意力权重的解释,说明模型为何认为某个时空模式是异常的。这不仅能增加工程师对AI决策的信任,还能为后续的工程诊断提供有价值的线索。

4 挑战与未来展望

尽管AI为大坝安全监测带来了革命性的机遇,但其全面工程化仍面临诸多挑战:(1)数据质量与治理:AI模型的性能上限由数据质量决定。如何系统性地解决传感器故障、数据缺失、噪声干扰等问题,建立高质量、标准化的监测数据库,是应用的前提。(2)模型泛化能力:在一座大坝上训练的模型,能否有效应用于地质条件、坝型、规模不同的另一座大坝?提升模型的跨坝型、跨地域泛化能力是亟待解决的科学问题。(3)实时性与边缘计算:对于需要秒级响应的突发险情预警,必须将部分AI模型部署到靠近数据源的边缘设备上,这对模型的轻量化和计算效率提出了更高要求。(4)人机协同与伦理规范:AI应定位为辅助决策工具,而非完全替代人类专家。如何设计高效的人机交互界面,以及制定AI在安全监测领域的伦理准则和责任边界,是社会层面必须回答的问题。展望未来,物理信息驱动的人工智能

(Physics-InformedAI)被视为最有前景的方向。该范式将大坝结构力学的基本方程(如平衡方程、本构关系)作为硬约束或软正则项嵌入到AI模型的训练过程中。这样构建的混合模型,既保留了数据驱动方法从海量数据中学习复杂模式的能力,又确保了其预测结果符合物理世界的客观规律,从而在保证高精度的同时,极大地提升了模型的鲁棒性、泛化能力和内在可解释性。这将是实现大坝安全监测从“被动响应”迈向“主动预见”和“智能自治”的关键一步。

5 结语

人工智能中的机器学习与深度学习正深刻改变大坝安全监测数据异常识别。本文梳理了从统计学习到深度学习的异常识别算法谱系,探讨了各方法机理与适用边界。强调算法需与大坝监测数据时空耦合、多源异构等特性深度融合,通过时空联合建模、多源信息融合等关键技术构建有效智能识别系统。虽面临数据质量、模型泛化等挑战,但以物理信息驱动为核心的新一代混合智能模型,为大坝安全智能预警指明方向。未来工程实践应打通“数据-算法-物理-决策”全链条,使人工智能成为守护江河、保障民生的智慧助力。

参考文献

- [1]刘毅,雒翔宇,赵宇飞,等.大坝工程智能建造技术与实践探索[J].中国水利,2025,(16):46-52.
- [2]龙耿文,袁树才,罗家良,等.水利工程大坝安全监测技术与发展研究[J].水上安全,2025,(20):148-150.
- [3]汪昌港,耿峻,赵鹏,等.基于模糊聚类和孤立森林算法的大坝监测数据粗差识别方法[J].水力发电,2025,51(12):96-100.
- [4]钱信超,赵小月,方伟.水电站大坝安全监测数据采集与分析方法研究[J].水上安全,2025,(08):142-144.