

基于多模态信息融合的工业生产线时序数据 异常识别与定位方法

徐辉

淮北职业技术学院

摘要: 面向工业生产线多源异构数据采集频率不一致、异常样本稀缺与定位解释不足等问题,提出一种基于多模态信息融合的时序数据异常识别与定位方法。首先构建统一时间栅格对齐机制,将传感器信号、视觉帧序列、声学时频特征与 PLC/SCADA 事件日志进行窗口化表征,并采用单模态编码器提取鲁棒特征。其次设计跨模态注意力融合网络,引入预测与重构联合自监督目标学习正常运行模式,同时以跨模态一致性偏差刻画协同破坏型异常。最后构建异常评分与可解释定位策略,将异常贡献分解到工位、模态与关键时间子段,实现“发现—定位—溯源”的闭环输出。实验结果表明,该方法在误报控制、早期预警与工位定位准确率方面优于单模态与简单融合基线,具备工程部署可行性。

关键词: 多模态融合; 时序异常检测; 自监督学习; 跨模态注意力; 工业生产线

随着智能制造与工业互联网的深入推进,工业生产线呈现出“多源感知、强耦合、强时序”的运行特征:振动、电流、温度、压力等传感器连续记录设备负载与健康状态,视觉与声学数据补充刻画工件外观、装配过程及摩擦碰撞、气路泄漏等早期征兆,PLC/SCADA 日志则反映工艺节拍与控制指令的执行轨迹。在此背景下,生产线异常不再仅表现为单一变量越界,而常以弱漂移、间歇突变与跨模态协同破坏等形式出现,并伴随噪声干扰、工况切换与数据缺失等现实因素,导致传统阈值规则、单模态预测或重构模型在误报控制、早期预警与可解释定位方面存在明显不足。尤其在异常样本稀缺、故障类型长尾分布的条件下,依赖大量标注的监督学习难以稳定覆盖现场需求;而仅输出异常分数的“黑盒报警”又难以支撑运维人员快速完成“哪一工位、哪一时段、哪一模态”触发异常的精准排查。基于此,本文聚焦多模态信息融合条件下的工业生产线时序异常识别与定位问题,构建统一时间对齐与窗口化表征,融合跨模态注意力与自监督学习以刻画正常运行模式,并通过异常贡献分解与关键时间段提取实现可解释定位,为预测性维护、质量监测与工艺优化提供兼具准确性与工程可用性的技术路径。

1 方法

1.1 问题定义与总体框架

设生产线在时间范围 $[1, T]$ 内采集到 M 种模态数据:连续时序传感器集合 $X_s = \{x_1, \dots, x_K\}$, 视觉帧序

列 $X_v = \{I_1, \dots, I_N\}$, 声学序列 $X_a = \{a_1, \dots, a_L\}$, 事件日志 $X_e = \{e_1, \dots, e_R\}$ 。目标是学习函数 F , 使其对每个时间窗 $W_\tau = [\tau - w + 1, \tau]$ 输出异常分数 $S_\tau \in [0, 1]$ 与定位结果 $L_\tau = \{\text{station}, \text{modal}, \text{segment}\}$ 。其中 station 为工位/设备单元, modal 为贡献最大的模态集合, segment 为触发异常的关键时间子段。

总体流程包括四步:1) 多模态采集与清洗;2) 统一时间栅格对齐与窗口化;3) 单模态编码器提取表示并进行跨模态融合;4) 异常评分与定位解释输出。为了兼顾实时性,推理阶段采用滑动窗口与增量更新,支持秒级报警与分钟级溯源。

1.2 多模态时间对齐与特征编码

工业数据常见“多频率+不规则事件”并存。本文引入统一时间栅格 Δt , 将连续传感器重采样到 Δt , 视觉与声学采用帧/块级聚合映射到同一栅格,事件日志则通过事件计数、事件类型嵌入与时间差特征转化为序列表示。为避免重采样造成关键信息丢失,连续信号在每个栅格内同时保留统计特征(均值、方差、峰峰值、峭度)与频域特征(主频能量、谱熵),并使用轻量 1D-CNN 提取局部形态;视觉模态采用 2D-CNN 或轻量 ViT 对关键区域特征进行编码,可结合产线 ROI 裁剪减少背景干扰;声学模态先计算 Mel 谱或 Log-Mel 特征,再用 2D-CNN 提取时频纹理;日志模态使用事件嵌入 +GRU 获得工艺状态表示。经过编码后,各模态在统一栅格上形成序列表示

基金项目: 2025AHGXZK30352安徽省重点科研项目融合多模态视角的工业时序数据异常检测模型研究; 2025AHGXSK50124安徽省特需专项项目双高校建设背景下旅游酒店产业学院“四维协同”育人模式创新与实践研究。

Zs,Zv,Za,Ze, 维度统一为 d 以便融合。

1.3 跨模态时序融合与自监督学习目标

融合模块采用跨模态 Transformer 实现“对齐 + 互补”。具体做法是将各模态序列拼接为多通道 token, 并加入两类位置编码: 时间位置编码 Pt 描述栅格顺序, 工位 / 模态编码 Pm 标识来源, 使模型能学习“同一时间不同模态的相关性”和“同一工位跨时间的演化规律”。注意力机制自动学习模态间权重, 当某一模态缺失或噪声增大时, 可由其他模态补偿, 从而提升鲁棒性。

在缺乏异常标签的条件下, 本文采用“预测 + 重构”的自监督联合目标刻画正常模式: 预测任务要求模型根据历史窗口预测未来 k 步表示, 约束其学习可持续的动态; 重构任务要求模型重构当前窗口的多模态表示, 捕捉局部细节一致性。联合损失为 $L = \lambda_1 \cdot L_{pred} + \lambda_2 \cdot L_{rec} + \lambda_3 \cdot L_{con}$, 其中 L_{con} 为跨模态对比学习项, 促使同一时间窗的不同模态表示在语义空间对齐、不同时间窗区分开, 减少对对齐误差对检测的影响。训练完成后, 模型形成对正常行为的紧致表示流形, 为异常检测提供基准。

1.4 异常评分与定位机制

异常评分由三部分组成: 预测残差 E_p 衡量未来偏离程度, 重构残差 E_r 衡量当前偏离程度, 跨模态一致性偏差 E_c 衡量模态间协同破坏程度, 最终 $S = \sigma(\alpha \cdot E_p + \beta \cdot E_r + \gamma \cdot E_c)$ 。其中 σ 为归一化函数, α, β, γ 可由验证集或在线自适应校准获得, 以平衡早期预警与误报率。

定位机制强调“可解释、可落地”。本文将异常分数分解到工位与模态: 在融合注意力矩阵中, 统计各工位 token 对异常得分的贡献, 得到工位贡献向量 $G_{station}$; 在模态维度上, 统计注意力权重与残差敏感性 (对输入特征做梯度归因或遮挡试验) 得到模态贡献向量 G_{modal} 。随后在时间维度上对残差序列做变点检测, 输出关键子段 segment, 实现“何时开始异常、持续多久、是否与节拍切换相关”的解释。最终输出不仅给出报警, 还给出“高贡献工位 + 高贡献模态 + 关键时间段”的定位三元组, 便于维修与工艺调整。

2 结果

2.1 实验场景与数据设置

为评估所提方法的工程适用性, 实验选取离散制造装配线作为对象, 覆盖上料、拧紧与在线检测等典型工位, 并同步采集连续传感器、视觉、声学 with PLC/SCADA 日志四类模态数据。连续信号包括电流、振动、温度、气压, 用于反映设备负载与结构状态; 视觉数

据聚焦关键工序区域以捕捉对位偏差、表面异常与画面污染; 声学数据用于捕获摩擦、碰撞与泄漏等早期征兆; 日志数据记录节拍与状态切换, 为异常溯源提供工艺语境。针对采样频率差异、网络延迟与短时缺失问题, 实验采用统一时间栅格完成对齐, 并以滑动窗口切分序列, 使模型在推理阶段可持续输出窗口级异常分数与定位线索。训练数据主要来自稳定生产时段以学习正常模式; 评估阶段引入夹具松动、扭矩漂移、相机镜头污染与气路微漏等典型异常片段, 兼顾突发性与渐变型异常。

2.2 异常识别效果与对比分析

对比结果显示, 单模态阈值规则在换型调参、节拍切换等场景下易误报, 对弱漂移异常预警滞后; 单模态预测残差在噪声较强或耦合复杂时易被非故障波动干扰; 单模态重构模型对视觉污染、外观异常较敏感, 但可能受光照与遮挡影响。简单拼接融合虽能利用多源信息, 却对齐误差敏感, 出现时钟漂移或丢帧时稳定性下降。

本文方法通过跨模态注意力实现动态加权融合, 并以“预测 + 重构”联合自监督目标刻画正常运行规律, 在误报控制与检测时效上更均衡。对于扭矩漂移等渐变异常, 模型可利用预测偏离实现更早预警; 对于夹具松动等突发异常, 多模态同步证据被融合强化, 从而降低噪声触发误报; 对于气路微漏等协同破坏型异常, 引入的一致性偏差能够捕捉模态间耦合关系变化, 减少漏检。

2.3 定位效果与机制验证

在定位层面, 方法输出工位、模态与关键时间子段的贡献解释, 便于现场排障。夹具松动通常表现为振动能量拾升并伴随视觉 ROI 轻微抖动, 注意力会聚焦夹紧相关工位并给出清晰变点时间段; 相机镜头污染则以视觉重构残差显著升高为主, 贡献分解可直接指向视觉模态, 避免无效检修。消融结果表明, 去除跨模态对比约束会削弱对齐并增加节拍切换误报; 弱化跨模态融合会降低互补能力; 移除一致性偏差项会削弱对耦合破坏型异常的敏感度并使定位更分散, 说明对齐、融合与一致性建模是性能提升的关键来源。

3 讨论

3.1 工程部署价值与应用边界

从工程角度看, 本文方法的价值不仅在于提升识别精度, 更在于把异常从“黑盒报警”转化为“可定位、可解释的证据链”。通过输出工位与模态贡献, 运维人员可优先检查高贡献单元, 缩短 MTTR; 通过关键时间子段, 工艺工程师可追溯异常与节拍、换线、参

数下发等事件的因果关联,形成持续改进闭环。同时,自监督训练主要依赖正常数据,适配异常稀缺的产线现实,且可通过在线增量方式定期更新,缓解设备老化与工艺漂移带来的分布变化。

但也需明确应用边界:当产线长期处于“非稳定工况”(频繁换型、频繁调参)时,正常模式本身多峰且快速变化,若缺乏工况标签或配方信息,模型可能将合法变化误判为异常;当关键模态长期缺失或传感器故障导致数据质量持续低劣时,融合模型虽有补偿能力,但定位解释会变得不可靠。因此在部署时应配套数据质量监控、工况标识接入与报警分级策略,将“数据异常”和“设备异常”区分开来。

3.2 方法局限与改进方向

本文仍存在三方面可进一步提升:第一,时间对齐在复杂网络延迟与设备时钟漂移下仍可能出现系统性偏差,后续可引入基于动态时间规整或可学习对齐模块,实现端到端软对齐;第二,定位解释目前主要依赖注意力与梯度归因,虽然直观,但在极端噪声下可能不稳定,未来可结合因果推断或故障树先验,把解释从相关性推进到更强的因果证据;第三,异常类型多样且跨产线迁移困难,后续可探索联邦学习与领域自适应,使不同工厂在不共享原始数据的前提下共享正常表征,提高冷启动效率。

4 结语

总体而言,基于多模态信息融合的时序异常识别与定位是智能制造落地的关键能力之一。本文提出的

“统一对齐+跨模态自监督融合+可解释定位”框架在识别精度、预警时效与定位可用性方面体现出综合优势,可为工业生产线的预测性维护、质量监测与工艺优化提供可复用的方法路径。

参考文献:

- [1] 陈锦涛,张逸,张良羽,等.基于用电数据挖掘的企业环保异常识别[J].电力建设,2025,46(02):74-87.
- [2] 刘强,卓洁,郎自强,等.数据驱动的工业过程运行监控与自优化研究展望[J].自动化学报,2018,44(11):1944-1956.DOI:10.16383/j.aas.2018.c180207.
- [3] 张成,戴絮年,李元.基于DPCA残差互异度的故障检测与诊断方法[J].自动化学报,2022,48(01):292-301.DOI:10.16383/j.aas.c190884.
- [4] 刘强,孔德志,郎自强.基于多级动态主元分析的电熔镁炉异常工况诊断[J].自动化学报,2021,47(11):2570-2577.DOI:10.16383/j.aas.c190313.
- [5] 沈潇军,葛亚男,沈志豪,等.一种基于LSTM自动编码器的工业系统异常检测方法[J].电信科学,2020,36(07):136-145.
- [6] 孙海丽,龙翔,韩兰胜,等.工业物联网异常检测技术综述[J].通信学报,2022,43(03):196-210.
- [7] 吴越,曹国彦.面向工业传感网络的时间序列异常检测综述[J].网络与信息安全学报,2024,10(04):17-36.
- [8] 尚书一,李宏佳,宋晨,等.互联网服务场景下基于机器学习的KPI异常检测综述[J].计算机研究与发展,2025,62(01):207-231.