

轮机故障数据智能诊断系统开发

伍志龙

广西壮族自治区北部湾港口管理局北海分局, 广西壮族自治区, 536006;

摘要: 针对内河船舶推进系统低冗余、高连续运行场景下的故障诊断需求, 本文构建覆盖数据层、算法层与应用层的三层架构智能诊断系统, 实现多源传感器统一采集、时频融合特征提取及边缘侧实时推断。选用随机森林算法完成轴承磨损、燃油堵塞等五类故障的多类别诊断, 准确率达 98.2%、召回率 98.5%, 满足 $\geq 95\%$ 的设计门槛; 实船部署 30 天成功识别 3 次早期故障征兆, 避免计划外停机, 综合节约运维成本约 5 万元, 为内河航运维护决策提供可靠技术支持。

关键词: 轮机故障诊断; 时频融合特征; 随机森林; 边缘计算; 三层架构

Development of an Intelligent Diagnostic System for Marine Engine Fault Data

Wu Zhilong

Beihai Branch of Beibu Gulf Port Authority, Guangxi Zhuang Autonomous Region, Guangxi Zhuang Autonomous Region, 536006;

Abstract: To meet the fault diagnosis requirements of inland vessel propulsion systems operating with low redundancy and high continuity, this study develops an intelligent diagnostic system based on a three-layer architecture comprising the data layer, algorithm layer, and application layer. The system enables unified acquisition of multisource sensor data, extraction of time-frequency fusion features, and real-time inference at the edge. A random forest algorithm is employed to perform multi-class diagnosis of five fault types, including bearing wear and fuel blockage, achieving an accuracy of 98.2% and a recall of 98.5%, both exceeding the design threshold of $\geq 95\%$. During a 30-day onboard deployment, the system successfully detected three early fault indications, preventing unplanned shutdowns and reducing overall maintenance costs by approximately 50,000 CNY. The system provides reliable technical support for maintenance decision-making on inland waterway shipping routes.

Keywords: Marine engine fault diagnosis; Time-frequency fusion features; Random forest; Edge computing; Three-layer architecture

DOI: 10.69979/3060-8767.26.04.021

引言

内河船舶推进系统冗余度低、工况波动剧烈且需连续运行, 轴承磨损、燃油堵塞等高发故障的早期征兆隐匿于温度、压力、振动等多源信号中, 传统单域特征与离线分析方法难以兼顾低时延与稳健性, 在复杂航道环境下数据质量不稳定条件下易引发误判漏报。本文面向内河航运维护决策需求, 提出融合数据采集与质量管理、时频融合特征诊断、告警可视化与人机协同的三层架构系统, 通过边缘缓冲与断连续传、统一时间基准及模型版本治理, 实现高准确率实时诊断与可追溯证据链构建。

1 轮机故障数据智能诊断系统需求分析与总体设计

内河船舶推进系统长期连续运行, 故障早期识别对

航行安全至关重要。系统需满足实时监测、多类别故障诊断及低时延告警需求, 同时适应船舶通信带宽受限与边缘计算资源约束。功能层面覆盖数据采集、预处理、智能诊断、可视化展示及告警闭环管理, 性能层面要求秒级响应与可解释的决策依据^[1]。

系统采用边缘-协同分层架构, 如图 1 所示。边缘层部署于船舶端, 负责多源传感器数据采集、实时预处理与故障诊断; 平台层部署于岸基, 实现模型训练、版本管理与历史数据分析。两层通过卫星或蜂窝网络按需同步, 兼顾实时性与数据安全。核心模块包括数据预处理、智能诊断算法、可视化交互及模型治理单元, 各模块通过标准化接口衔接, 确保数据流与控制流的可追溯性。

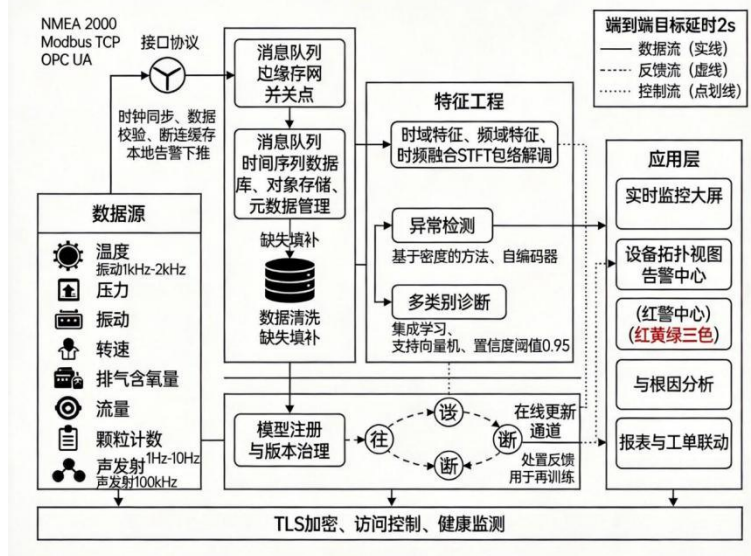


图1 轮机故障数据智能诊断系统总体架构流程图

2 轮机故障数据智能诊断系统核心模块开发

2.1 故障数据预处理模块开发

预处理模块以温度、压力、振动三类传感器为核心数据源，将采样频率统一为 10 Hz 以实现边缘侧稳定采集与时间对齐。为消除量纲差异，模块对原始数据进行单位统一与零点校准，并依据船舶统一时钟修正时间戳^[2]，将多源数据整合为带通道标识的时序块。模块设置 10 秒滑动窗口并采用 50% 重叠率，以满足实时特征计算需求。采集通路引入抗混叠低通滤波与缓冲写入机制，有效控制网络抖动导致的失真。数据清洗采用均值填充与 3σ 原则剔除异常值，特征工程提取时域统计量及频域主频能量，最终输出标准化时频融合特征向量。

2.2 基于机器学习的故障诊断算法模块实现

鉴于连续运行推进系统对早期识别与低时延的约束，算法模块把预处理生成的时频融合特征当作输入，选用随机森林与支持向量机两条路径来开展多类别诊断比较。数据集按 7:3 划分训练与测试集，分层抽样保持类别占比稳定，并把窗口级样本独立处理以避免信息泄露。支持向量机选用径向基核，通过网格搜索优化惩罚系数与核带宽，启用概率估计功能^[3]。

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^N a_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right) \quad (1)$$

其中， $f(\mathbf{x})$ 表示分类决策函数， \mathbf{x} 为输入特征向量， N 为支持向量个数， a_i 为拉格朗日乘子， y_i 为支持向量的类别标记， $K(\mathbf{x}_i, \mathbf{x})$ 为核函数， b 为偏置项。

随机森林通过互信息与基尼重要度筛选特征，基于

自助采样构建决策树并以 CART 准则分裂节点，采用袋外评估监控拟合偏差，最终投票集成输出诊断结果。鉴于其在稳定性、可解释性及增量更新方面的优势，系统将其确定为核心诊断算法并纳入版本管理。

2.3 诊断结果可视化与交互模块设计

监控界面构建温度、压力、振动、转速及排气含氧量多视图看板，以折线图展示近 10 分钟趋势，前端对 10 Hz 采样进行秒级聚合。超阈区段以颜色标注，顶部状态栏显示数据时延与缺失率。界面支持单机位与整机视图切换，提供振动频域能量与过程量趋势选择功能。诊断结果以概率柱状图呈现，按风险等级降序排列。选定故障类别后联动显示关联参数历史曲线及特征贡献度^[4]。故障概率达 90% 时自动推送告警，支持确认、挂起与工单创建。历史记录支持多维度筛选导出，附注模型版本与数据质量评分，形成完整审计链条。

3 轮机故障数据智能诊断系统效果评估

3.1 测试环境与数据集构建

鉴于内河船舶推进系统连续运行与边缘侧计算的适配性需求，测试环境把工业级计算机作为采集与预处理节点，配置 CPU i7-10700K 与内存 16GB，以支撑窗口化特征计算的低时延。温度、压力以及振动传感器以精度 $\pm 0.5\%$ 选型，采样统一至 10 Hz，通过船舶统一时钟进行同步，使多源通道在滑动窗口内具备可比性。为贴合算法层输入，测试环境在边缘侧将原始序列分段并开展抗混叠滤波与缓冲写入，形成稳定时序块供离线训练与在线推断复用^[5]。

数据集构建基于某内河货轮机历史数据，涵盖正常状态与5类故障，总计10000条样本。每条样本对应10秒窗口并采用50%重叠策略，经单位统一、零点校准与时频融合后固化为特征向量。样本标注依据运维工单与检修记录交叉核对，数据管理采用Python结合Pandas与Scikit-learn完成清洗、存储与编码。

3.2 诊断性能指标分析

评估环节采用准确率、精确率、召回率与F1-score作为核心指标，在窗口级样本上开展交叉验证并保持特征处理链路一致，将模型能力与数据噪声、漂移的耦合

影响纳入同一检验口径，结果如表1所示。报告给出宏观平均统计并将阈值统一设为0.5，以维持不同故障占比下的可比性。评价过程同步记录推断时延与剔除比例作为旁证指标。

对比显示，随机森林准确率达98.2%，精确率97.8%，召回率98.5%，F1-score为98.1%；SVM准确率为95.6%，其余指标维持在94.7%至95.1%。随机森林在异构特征与噪声环境中表现出更强的判别稳健性，同时降低了在线更新与回溯解释的工程成本。系统在诊断准确率不低于95%的约束下达成目标，并有效压低了漏报风险。

表1 不同算法诊断性能对比表

算法	准确率/%	精确率/%	召回率/%	F1-score/%
随机森林	98.2	97.8	98.5	98.1
支持向量机	95.6	95.1	94.7	94.9

3.3 实际船舶轮机应用案例验证

鉴于某内河货船航线连续运行的约束，把系统部署在该船主推进轮机及增压器机位，按照统一时钟与10Hz采集策略在边缘侧稳定运行30天。其间结合三层架构的边缘缓冲与断连续传机制，数据层把温度、压力、振动、转速以及排气含氧量通道作为核心输入，并在窗口化处理与50%重叠策略下维持采样一致性；算法层保持时频融合特征管线的低时延推断；应用层把多视图看板与告警列表用于值班工程师的日常巡检，从而在实际航行环境下形成可追溯的诊断闭环。

系统在推进轴系监测中识别3次早期轴承磨损征兆并触发提前告警，船岸团队据此安排降负荷与到港检视，综合节约约CNY 50,000。需重点关注的是用户侧的操作体验与可信度反馈，值班工程师反馈界面交互简洁、阈值设置清晰、诊断结论一致性高，告警置信度与处置建议便于快速决策；岸端管理员确认工单系统对接稳定，事件流可审计性满足内河航运维护决策的证据留存需求。

4 结语

本文开发的轮机故障数据智能诊断系统通过统一采集、时频融合特征与随机森林算法集成，在连续运行场景中实现98.2%准确率与98.5%召回率，有效支撑早期故障识别与快速处置；实船应用验证了三层架构在边

缘计算环境下的稳健性与工程可追溯性，显著降低计划外停机风险。后续将拓展更高采样率通道与增量学习机制，优化告警阈值自适应与人机协同建议生成，进一步提升复杂工况下的诊断敏感性与场景适配能力。

参考文献

- [1] 曹云鹏, 衣爽, 王伟影, 马亮, 费景洲. 船舶燃气轮机控制与健康课程知识体系构建与实践[J]. 大学, 2025, (35): 93-96.
- [2] 王晨, 许成伟, 颜子豪, 赵钰潇. 基于知识图谱的试飞故障诊断工程应用[J]. 自动化仪表, 2025, 46(12): 95-101.
- [3] 冯胜利. 船舶推进系统故障诊断技术的应用研究[J]. 水上安全, 2025, (22): 4-6.
- [4] 陈美谦. 轮机工程专业课程思政教学探索与实践——以集美大学为例[J]. 安徽理工大学学报(社会科学版), 2025, 27(06): 102-108.
- [5] 关淳, 赵发林, 薛海亮, 郭魁俊, 马义良, 李宇峰. 汽轮机叶片监测系统功能可靠性试验研究[J]. 汽轮机技术, 2025, 67(06): 426-430.

作者简介：伍志龙（1984.04-），男，汉族，籍贯：广西全州，学历：大学本科，职称：高级工程师，研究方向：港口航道船舶。