

# 大模型工业设备智能运维系统设计

刘 智

连云港市花果山风景区管理处 江苏连云港 222061

**摘要：**在工业4.0战略持续推进的背景下，工业设备呈现出大型化、复杂化与智能化的发展趋势，传统依赖人工经验的运维模式已无法满足设备高效稳定运行的需求。本文提出并设计了一套基于大模型的工业设备智能运维系统，通过整合多源数据采集、智能预处理、大模型深度分析、故障预警及运维决策优化等核心功能，实现对工业设备全生命周期的智能化管理。该系统以大模型为核心驱动，依托其强大的数据分析与逻辑推理能力，显著提升设备故障诊断的准确性与前瞻性，有效降低运维成本，提高设备运行效率。论文详细阐述了系统的设计背景、总体架构、关键技术实现及各功能模块设计方案，并对系统未来发展方向进行展望，为工业设备智能运维领域的理论研究与实践应用提供参考。

**关键词：**大模型；工业设备；智能运维；系统架构；故障预警；数据安全

## 一、系统设计原则与目标

### （一）设计原则

为确保系统的科学性、实用性与可靠性，设计过程中严格遵循以下原则：

**实用性原则：**紧密结合工业设备运维实际需求，功能模块设置与操作流程设计符合运维人员工作习惯，确保系统可快速落地应用，解决实际运维问题。

**可靠性原则：**采用高可靠硬件设备与冗余软件架构，具备较强的容错能力与抗干扰能力，保障在复杂工业环境下数据采集、传输、处理的准确性与完整性。

**可扩展性原则：**采用模块化架构设计，各功能模块独立解耦，支持功能扩展与版本升级；同时兼容多种类型工业设备数据接入，满足不同场景运维需求<sup>[1]</sup>。

**实时性原则：**优化数据处理流程与故障响应机制，确保设备异常状态可及时识别，预警信息与诊断结果可快速反馈，满足工业生产对运维实时性的要求。

**安全性原则：**构建全流程数据安全防护体系，通过数据加密、访问控制、身份认证等技术，防止数据泄露、篡改与非法访问，保障企业核心数据安全。

### （二）设计目标

系统的核心目标是构建一套集数据采集、预处理、

大模型分析、故障预警、运维决策于一体的全流程智能运维系统，具体目标包括：

**数据采集全面化：**实现设备多源异构数据（运行参数、振动数据、温度数据、故障记录等）的全面采集，采集频率与精度满足大模型分析与故障诊断需求。

**故障预警精准化：**基于大模型分析，实现设备潜在故障的提前5-7天预测，故障预警准确率不低于90%，降低误报与漏报率。

**故障诊断高效化：**设备故障发生后，10分钟内完成故障定位与原因分析，自动生成诊断报告与维修建议。

**运维决策智能化：**结合设备状态、故障信息与维修资源，制定最优运维方案，优化维修时间、人员与备件调配，提升运维效率。

**系统运行稳定化：**系统平均无故障运行时间不低于1000小时，数据传输成功率不低于99.9%，保障长期稳定运行。

## 二、系统总体架构设计

基于大模型的工业设备智能运维系统采用分层架构设计，自下而上分为感知层、传输层、数据层、大模型分析层与应用层，各层协同配合，共同完成设备智能运维任务。

### （一）感知层

感知层是系统数据采集的基础，负责获取设备运行状态数据与环境数据，主要包括三类采集设备：

**状态监测传感器：**如振动传感器、温度传感器、压

**作者简介：**刘智（1985.05--），男，汉族，江苏连云港人，职务/职称：工程师（电子信息工程），学历：大学本科，研究方向：人工智能、智慧景区、AI大模型。

力传感器、转速传感器等，实时采集设备振动幅值、温度变化、压力值、转速等核心运行参数，直接反映设备运行状态<sup>[2]</sup>。

过程数据采集设备：如PLC、DCS系统，采集设备生产过程数据（产量、能耗、物料消耗等），为设备运行效率与生产效益分析提供数据支撑。

环境监测传感器：如湿度传感器、粉尘传感器、噪音传感器，采集设备运行环境的湿度、粉尘浓度、噪音分贝等数据，分析环境因素对设备运行的影响。

感知层采用分布式部署方式，设备直接安装于工业设备本体或周边区域，将采集的模拟信号转换为标准数字信号，为后续数据传输与处理奠定基础。

## （二）传输层

传输层承担感知层与数据层之间的数据传输任务，采用“有线+无线”混合传输模式，满足工业场景下数据传输的实时性、可靠性与安全性要求：

有线传输：针对数据量大、传输距离近且实时性要求高的场景（如车间内设备），采用工业以太网（Profinet、EtherNet/IP）传输数据，具备传输速率高、稳定性强、抗干扰能力突出的优势。

无线传输：针对设备分布分散、布线困难或移动设备（如AGV小车），采用LoRa、NB-IoT、5G等工业无线技术。其中，LoRa适用于远距离、低功耗、小数据量传输；NB-IoT支持广覆盖、大连接；5G则满足高带宽、低时延、大连接需求。

传输层同时部署数据安全机制，采用SSL/TLS协议对传输数据加密，通过防火墙与访问控制策略抵御网络攻击，保障数据传输安全。

## （三）数据层

数据层负责数据的存储、管理与预处理，为大模型分析层提供高质量数据支持，主要包括数据存储与数据预处理两大模块：

### 1. 数据存储模块

针对工业设备运维数据量大、类型多样、时效性强的特点，采用“关系型数据库+非关系型数据库+时序数据库”混合存储架构：

关系型数据库（MySQL、Oracle）：存储结构化数据，如设备基本信息（编号、型号、厂家）、运维人员信息、维修记录等，保障数据一致性与查询效率<sup>[3]</sup>。

非关系型数据库（MongoDB、HBase）：存储非结构化与半结构化数据，如故障图片、维修视频、传感器原

始数据文件，具备存储容量大、扩展性好的优势。

时序数据库（InfluxDB、TimescaleDB）：存储设备时序数据（振动、温度等），针对时序数据存储与查询优化，支持海量数据高效读写与趋势分析。

### 2. 数据预处理模块

针对感知层采集数据中的噪声、缺失值、异常值等问题，通过以下流程优化数据质量：数据清洗：采用移动平均滤波、卡尔曼滤波去除噪声；通过线性插值、多项式插值填补缺失值；利用 $3\sigma$ 准则、箱型图法识别并剔除异常值。

数据集成：整合多源数据，消除冗余与不一致性，形成统一数据视图。

数据转换：对数据进行标准化或归一化处理，将参数转换为[0, 1]区间数值，避免数据量级差异影响模型分析结果。

数据降维：采用PCA、LDA等算法降低高维数据维度，减少计算复杂度，提升处理效率。

## （四）大模型分析层

大模型分析层是系统核心，负责数据深度分析与智能决策，包含四大模块：

### 1. 大模型训练模块

模型选型：结合设备数据特点，选择适配模型架构。如基于Transformer的TFT模型处理时序数据，捕捉长期依赖关系；结合知识图谱的模型整合领域知识，提升可解释性。

数据集构建：从数据层选取历史数据，划分训练集、验证集与测试集，分别用于模型参数学习、超参数调整与性能评估。

模型训练：采用TensorFlow、PyTorch分布式框架训练模型，通过调整学习率、batchsize等超参数优化性能，利用正则化、dropout技术防止过拟合。

模型评估与更新：通过准确率、召回率、RMSE等指标评估模型性能，达标后部署应用；定期利用新数据微调模型，确保适配设备状态变化。

### 2. 故障预警模块

数据输入：接收数据层预处理后的实时数据。

状态预测：大模型分析数据，预测未来5-7天设备运行参数趋势，判断故障风险。

预警判断：对比预测值与正常阈值，划分低、中、高风险等级，触发对应预警。

预警通知：通过短信、邮件、系统弹窗推送预警信

息，展示异常参数、风险等级与预测故障时间。

### 3. 故障诊断模块

故障特征提取：从实时与历史数据中提取故障特征（如振动频谱、温度趋势）。

故障分类识别：大模型分析特征，识别故障类型（如轴承磨损、电机过载）。

故障原因分析：结合历史数据与领域知识，分析故障诱因（如超负荷运行、备件老化）。诊断报告生成：自动生成报告，包含故障设备、类型、时间、原因及维修建议<sup>[4]</sup>。

### 4. 运维决策模块

维修优先级排序：根据设备重要性、故障影响范围、生产紧急程度排序维修任务。

维修资源分配：匹配维修人员技能、调配备件与设备，避免资源短缺延误维修。

维修计划制定：结合生产计划确定维修时间与生产调整方案，减少对生产的影响。

运维效果评估：维修后分析设备状态恢复情况与成本收益，优化后续方案。

## （五）应用层

应用层是系统与用户交互的界面，为不同角色用户提供个性化服务，主要包括：

设备状态监测模块：以图表（折线图、仪表盘）实时展示设备参数，支持历史数据查询与回放。

故障预警与诊断模块：展示预警信息与诊断报告，支持运维人员处理预警、查看维修建议。

运维管理模块：管理维修任务、人员绩效与备件库存，跟踪维修进度。

报表统计与分析模块：自动生成设备运行效率、故障统计、运维成本报表，辅助管理决策。

系统管理模块：负责用户管理、权限分配与参数配置，保障系统安全运行。

应用层采用B/S架构，用户通过浏览器访问系统，实现跨平台、跨设备使用。

## 三、系统关键技术实现

### （一）大模型优化技术

模型轻量化：采用剪枝、量化、蒸馏技术减少模型参数与计算量。如剪枝去除冗余神经元，量化将32位浮点数转换为16位/8位整数，实现模型在边缘设备的部署。

领域知识融入：构建工业设备运维知识图谱，将设

备结构、故障树等知识融入模型训练，提升诊断准确性与可解释性。

增量学习：利用新数据更新模型，无需重新训练，适应设备状态变化与新故障类型，减少资源消耗。

### （二）数据安全与隐私保护技术

数据加密：存储采用AES加密敏感数据，传输采用SSL/TLS加密，防止数据泄露。

访问控制：基于RBAC模型分配权限，结合多因素认证，确保仅授权用户访问数据。

数据脱敏：对共享数据采用替换、屏蔽、泛化处理，去除敏感信息（如设备编号、企业名称）。

数据备份与恢复：本地与异地双重备份数据，制定恢复预案，保障数据可恢复性。

### （三）实时数据处理技术

边缘计算：在现场部署边缘节点，完成数据清洗、滤波等预处理，减少传输延迟。

流处理技术：采用ApacheFlink、KafkaStreams框架实时处理数据流，低延迟响应设备异常。

分布式计算：利用Hadoop、Spark框架并行处理海量数据，提升大模型训练与数据分析效率。

## 四、系统挑战与展望

### （一）面临挑战

大模型与工业场景适配性：通用大模型在处理工业时序数据、噪声数据时性能不稳定，多设备类型通用模型构建难度大。

数据质量与数量：工业现场数据存在噪声、缺失问题，新型设备与罕见故障标注数据不足，影响模型性能。

实时性与资源平衡：大模型计算复杂度高，对硬件资源需求大，难以在保证实时性的同时降低资源消耗。

跨企业数据共享：企业因数据安全顾虑不愿共享数据，导致数据孤岛，制约模型泛化能力提升。

### （二）未来展望

大模型专业化与轻量化：开发行业专用（如汽车制造、钢铁）与设备专用大模型，优化轻量化技术，实现边缘部署<sup>[5]</sup>。

数据质量与增强：升级传感器与采集设备，优化预处理算法；采用时序数据扩充、合成故障数据技术，补充标注数据。

边缘-云端协同：边缘节点处理实时任务，云端承担复杂模型训练，平衡实时性与资源消耗。

隐私计算与数据共享：引入联邦学习、安全多方计

算技术，在保护隐私的前提下实现跨企业数据联合训练，推动行业协同发展。

### 结论

本文设计的基于大模型的工业设备智能运维系统，通过分层架构实现数据采集、处理、分析与应用的全流程智能化，以大模型为核心提升故障预警与诊断能力，结合关键技术保障系统可靠性、安全性与实时性。系统可有效降低运维成本、减少停机时间，满足工业设备智能运维需求。未来，随着大模型与工业技术的深度融合，系统将进一步优化，为工业行业数字化转型提供更强支撑。

### 参考文献

- [1] 刘哲, 姜永利. 变电站智能化运维系统的设计与实现研究[J]. 中国设备工程, 2025, (20): 28-30.
- [2] 李孟. 基于智能传感与机器人技术的核电设备远程运维系统研究[J]. 中国机械, 2025, (25): 122-125.
- [3] 肖楚生. 水电站机电设备智能运维系统的关键技术与应用展望[J]. 小水电, 2025, (04): 64-69.
- [4] 邓永祁, 杨将, 阳亦斌. 物联网时序数据库在城轨信号智能运维系统中的应用[J]. 控制与信息技术, 2025, (04): 92-98.
- [5] 史正强. 锐捷交换机设备的故障诊断与智能运维系统构建[J]. 信息记录材料, 2025, 26(08): 139-141.