

基于传感器技术和 I-LSTM 算法的风电机设备运行故障检测及诊断研究

孙 晔¹, 郭 琳²

(1. 秦皇岛港股份有限公司 第九港务分公司, 河北 秦皇岛 066000;

2. 秦皇岛港股份有限公司 第六港务分公司, 河北 秦皇岛 066000)

摘要: 有效的故障检测与诊断将极大地提高风电机设备运行效率和可靠性, 降低维修成本, 保障生产过程的顺利进行; 为实现高效率的设备故障预警与维护, 研究基于传感器技术和机器学习的设备运行故障检测及诊断方法; 采用箱型图法和小波包降噪法等对传感器传输的数据信号进行预处理; 使用双向长短时记忆网络构建时间序列预测模型; 并基于预测残差和贝叶斯概率理论, 设计信号异常识别策略, 对故障进行实时监测与故障预警; 经实验测试, 研究设计模型的诊断准确率为 98.88%, 无漏诊情况, 误诊率在 1.5% 以下, 实现了在提前 14 小时以上进行预警; 经实际应用, 研究设计模型满足了风电机设备故障预警的及时需求, 同时能够在较高的准确率下对故障进行诊断。

关键词: 传感器; 机器学习; 机械设备; 故障检测; 时序预测

Wind Turbine Equipment Operation Fault Detection and Diagnosis Research Based on Sensor Technology and I-LSTM Algorithm

SUN Ye¹, GUO Lin²

(1. No. 9 Port Branch of Qinhuangdao Port Co., Ltd. Qinhuangdao 066000, China;

2. No. 6 Port Branch of Qinhuangdao Port Co., Ltd. Qinhuangdao 066000, China)

Abstract: Effective fault detection and diagnosis will greatly improve the operational efficiency and reliability of wind turbine equipment, reduce maintenance costs, and ensure the smooth progress of production process. To achieve efficient equipment fault warning and maintenance, an equipment operation fault detection and diagnosis method based on sensor technology and machine learning is researched. the box plots and wavelet packet denoising methods are used to preprocess the data signals transmitted by sensors. the bidirectional long short-term memory network is used to construct the time series prediction model. Based on prediction residuals and Bayesian probability theory, a signal anomaly recognition strategy is designed to monitor and warn faults in real-time. Through experimental testing, the diagnostic accuracy of the research and design model is 98.88%, with no missed diagnosis and a misdiagnosis rate of below 1.5%, achieving early warning more than 14 hours in advance. Through practical application, the research and design model meets the timely needs of wind turbine equipment fault warning, and can diagnose faults with high accuracy.

Keywords: sensors; machine learning; mechanical equipment; fault detection; time series prediction

0 引言

随着工业社会的发展以及科学技术的不断进步, 现代工业开始全面使用机械设备代替人工。但在工作效率提高的同时, 机械设备的运行维护至关重要。若在工业生产过程中出现机械故障, 轻则影响生产效率, 重则引起安全事故^[1]。机械设备出现故障一般表现在无法完成特定的功能需求任务, 由于机械设备中存在较多复杂的零件, 机械的故障识别也存在一定的挑战^[2]。随着科学技术的不断发展, 目前机械设备的故障诊断技术也逐渐开始成熟, 机器学习

的使用成为研究重点内容。

但目前的智能诊断技术仍存在一定的局限性, 研究人员针对目前存在的问题, 开始探索诊断技术的优化策略^[3]。文献 [4] 针对道岔故障检测难以及时效性差等问题, 利用深度卷积神经网络和支持向量机构建了一种新的智能故障诊断方法。根据道岔转换曲线与标准曲线的相似度来识别道岔转换正常与异常情况。然后再采用时域特征提取和筛选方法, 利用智能算法进行故障诊断识别。实验表明, 该方法的故障诊断方法正确率达到了 99.01%^[4]。

收稿日期: 2024-02-19; 修回日期: 2024-04-05。

基金项目: 河北省科技计划项目(2015ZC20809)。

作者简介: 孙 晔(1979-), 男, 硕士, 高级工程师。

通讯作者: 郭 琳(1982-), 女, 大学本科。

引用格式: 孙 晔, 郭 琳. 基于传感器技术和 I-LSTM 算法的风电机设备运行故障检测及诊断研究[J]. 计算机测量与控制, 2024, 32(9): 51-57.

文献 [5] 为提高机械设备故障诊断的精度, 提出了一种分数阶域多尺度特征卷积神经网络的智能诊断方法。首先对原始振动信号进行分数傅里叶变换, 获得时频特征, 然后使用具有多尺度特征学习模块的卷积神经网络对故障信息进行识别。结果表明, 该方法能够实现较为理想的故障诊断^[5]。文献 [6] 针对轴承故障诊断过程中存在缺乏足量故障样本、信号分布差异大等问题, 提出一种基于改进联合分布适应的轴承智能故障诊断方法。首先利用小波包变换对振动信号进行分解与重构, 然后使用改进联合分布适应方法对特征级进行分布适应处理。最后利用源域特征样本对诊断模型进行训练。结果表明, 该方法的最高故障诊断准确率达到 100%^[6]。若仅仅只是对出现故障的设备进行检测诊断, 这只能是在设备出现故障之后及时止损, 并不能较大程度减少故障出现频率。而对设备运行故障进行实时诊断和预测能够进一步减少维护成本, 提高工业生产效率^[7]。为此, 研究利用传感器技术与机器学习对机械设备运行状态进行监测并预测故障, 在发现异常时进行智能预警。而设备运行过程中产生的时间序列能够显著代表设备的运行状态, 故利用设备运行过程中产生的时序信号能够有效预测识别信号异常状况。通过时间序列预测设备运行过程中可能出现的故障能够及时预警, 避免设备出现故障导致的事故^[8]。文献 [9] 为有效预测高铁列车车载设备故障率, 减少故障发生率, 提出一种基于变分模态分解和门控循环单元的故障率预测模型。其通过对设备的时间序列数据进行训练学习, 对序列变化情况进行预测。实验结果表明, 该方法得到的均方根误差以平均绝对误差分别为 0.045 和 0.039^[9]。文献 [10] 提出了一种基于门控循环单元的机械设备故障预测方法。首先对设备运行过程中产生的时序信号进行收集, 然后使用 CRITIC 权重法确定影响设备故障发生的关键参数, 最后根据历史数据进行预测分析^[10]。文献 [11] 为提高列车设备的现场检修效率, 提出了一种基于故障时序特征模式识别的实时诊断方法。首先通过机理分析选择与故障关联的指标, 然后利用高斯混合模型和隐马尔可夫链算法建立故障诊断模型。实验结果表明, 该方法能够有效实现设备故障的实时诊断^[11]。文献 [12] 针对织物防弹装置的响应以及故障模式进行了研究, 其采用多尺度数值模拟的方法分析了接触力、弹丸速度以及织物能量的演变。通过故障模式的分析及诊断, 对防弹装置记忆性优化^[12]。文献 [13] 针对涡轮叶片的失效问题, 对其失效机理以及故障预防方法进行了全面综述。其发现涡轮叶片和轮毂存在裂纹、疲劳、腐蚀、冲蚀、材料缺陷、热机械失效等多种失效类型。燃气轮机的失效主要有裂纹、腐蚀、疲劳失效等, 水轮机的失效主要有侵蚀和疲劳失效^[13]。

综合国内外研究现状可知, 目前的基于信号预测结果进行异常信号识别方法虽然可能能够达到较高的分析精度, 但并不能保证精度的稳定性, 在数据随机性过强的情况下, 分析精度依然较低, 信号异常情况划分也不够明确。为此,

研究将贝叶斯理论引入到信号异常识别中, 进一步捕捉预测模型得到的残差数据中的特征变化。由此, 研究利用物理型传感器采集机械设备上的物理信号, 并将其转换为电信号, 最后通过计算机转换为数字时序信号。然后利用改进的长短时记忆网络对信号进行预测, 根据预测残差对风电机机械设备进行实时的故障诊断。该模型旨在保证工业生产过程中的生产效率, 维护生产安全。

1 风电机机械设备运行故障诊断模型构建

1.1 基于传感器技术的数据采集及预处理

直驱式风电机机械设备在运行过程中会产生各种信号, 这些信号代表着其电压电流变化或振动幅度等等。若机械设备出现故障, 这些信号也会出现对应的异常状况, 对异常信号进行识别就能够检测设备是否存在故障^[14-16]。为采集到机械设备运行过程中产生的信号, 研究在机械设备上安装不同传感器对运行数据进行采集。但由于设备存在类型较多, 数据类型也很多, 还有很多数据存在缺失值, 冗余值, 异常值等情况, 这些数据并不能直接输入机器学习模型中。为此, 研究对信号数据进行预处理, 研究设计的预处理方法流程如图 1 所示。

采集数据中有许多异常值, 这些异常值的存在很大程度上影响了后续时间序列预测的准确性, 为此需要对异常值进行剔除。但使用肉眼进行异常值判断并剔除太浪费时间了, 为此, 研究利用箱型图法对离群异常数据进行识别, 将这些数据筛选出来。研究发现, 如果简单将这些异常数据进行剔除, 破坏时间序列的连续性, 为此, 研究将筛选出来的异常值统一按照缺失值进行处理。缺失值处理过程如式 (1) 所示:

$$x_i = \frac{x_{i-1} + x_{i+1}}{2}, \text{if } x_i \notin [D, U] \quad (1)$$

式 (1) 中, $[D, U]$ 为箱型图的上下限取值范围, x_i 为时间点 i 的观测值。在信号采集中, 会产生许多噪声, 若不对数据进行降噪处理, 数据中的有效信息将不足以后续研究使用。为此, 需要提高数据信噪比, 进而才能获得更加有效的分析结果。目前常用的提高信噪比的方法有平滑去噪和信号变换。两种方法相比来说, 信号变换降噪法能够在去噪的同时更好地保留数据的原有特征。因此, 研究使用信号变换降噪法中的小波变换对数据进行降噪处理。研究首先对信号进行分解, 在分解之后求得各个节点的小波系数。噪声过高的数据得到的小波系数将超过阈值, 故可根据计算得到的小波系数对部分数据进行过滤。过滤之后将数据信号按式 (2) 进行重构。

$$d_i^{j+1, n} = \sum_{k \in \mathbb{Z}} [h_{i-2k} d_k^{j, 2n} + g_{i-2k} d_k^{j, 2n+1}] \quad (2)$$

式 (2) 中, k 为继续分解的层数, \mathbb{Z} 为所有整数的集合, j 为频率指数, $d_i^{j+1, n}$ 为信号在分辨率为 $j+1$ 时的正交小波包分解系数, n 为离散数据点个数, 为 h_{i-2k} 和 g_{i-2k} 是由滤波器系数组成的序列。在确定阈值时, 研究选择 Bayes Shrink 小波阈值估计法进行阈值估计。计算方法如式 (3) 所示:

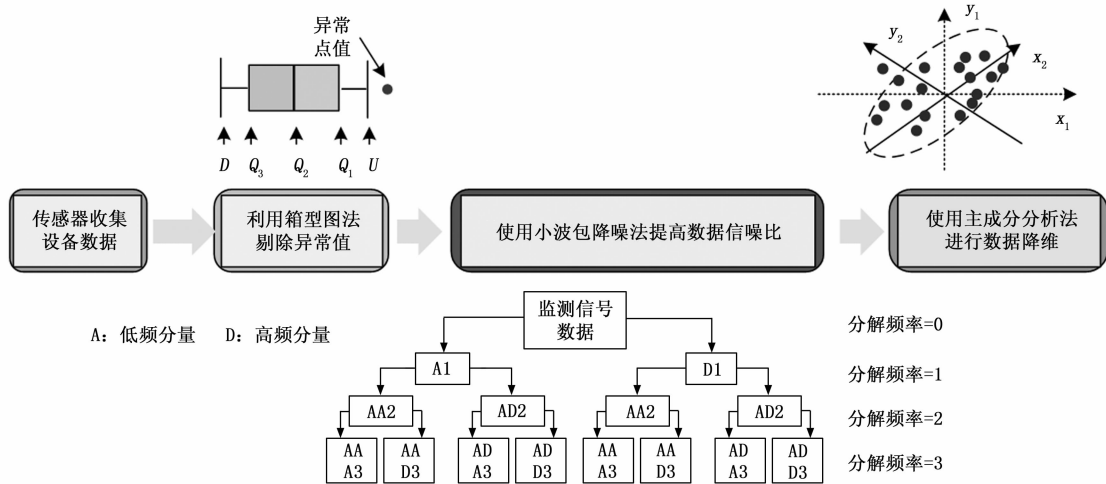


图 1 传感器采集数据的预处理流程

$$T(j, n) = \sigma^2(j, n) / \sigma_x(j, n) \quad (3)$$

式 (3) 中, σ_x 为原始信号方差, σ^2 是噪音方差, T 为所得阈值。通过软阈值函数对小波包系数进行过滤, 过滤过程如式 (4) 所示:

$$d_i^{j,n} = \begin{cases} \text{sgn}(d_i^{j,n}) (|d_i^{j,n} - T(j, n)|), & |d_i^{j,n}| \geq T(j, n) \\ 0, & |d_i^{j,n}| < T(j, n) \end{cases} \quad (4)$$

由于机械设备存在多样性, 采集到的数据存在不同的量纲, 而机器学习需要统一量纲进行训练。为此, 研究对数据进行消除量纲操作, 标准化的计算方法如式 (5) 所示:

$$\hat{x} = \frac{x - E(x)}{\sqrt{\text{var}(x)}} \quad (5)$$

式 (5) 中, x 和 \hat{x} 分别为处理前后的值, $E(x)$ 为样本均值。研究过程中发现, 小波变换在处理某些具有复杂特征和非平稳性的信号时, 具有一定的局限性, 如分解层数的选择、小波基函数的选取以及阈值处理的方式等。为此, 研究引入基于信号复杂度的自适应分解层数确定方法。通过分析信号的频率特性和能量分布, 动态地确定小波变换的分解层数, 从而能够在保证信号特征提取的同时, 避免过度分解导致的计算复杂度和信息冗余。针对阈值处理方式, 研究引入改进 Bayes Shrink 算法的阈值估计方法对小波变换进行优化改进。该方法在 Bayes Shrink 算法基础上, 加入对信号局部特征的分析, 从而能够更准确地估计小波系数的阈值。

为准确探测设备的故障信息, 确定故障位置, 研究在设备的同一个位置设备多个测点。但这也会导致收集信号存在冗余的情况。为避免过多的重复信息导致的分析结果不准确问题, 研究使用主成分分析 (PCA, principal component analysis) 对数据进行降维。但在研究过程中发现, 传统的 PCA 方法仍存在一定的局限性, 其简单丢弃非主成分数据, 但这些非主成分信息中可能还含有许多有用信息。为此, 研究将贝叶斯概率理论引入到降维过程中对 PCA 进行改进。改进后的 PCA 会在得到信号数据排序之后, 对有

用信息进行进一步分析和评估, 然后再确定是否需要丢弃。在降维过程中, 首先对特征信号的贡献率进行计算, 计算方法如式 (6) 所示:

$$v_i = \sum_{j=1}^k \frac{\lambda_j \eta_{ij}^2}{\sigma_{ii}} \quad (6)$$

式 (6) 中, v_i 为前 k 个成分对原变量的贡献率, σ_{ii} 为方差值, λ_j 为利用协方差矩阵分解特征后得到的值, η_{ij} 为线性变换。在计算得到各特征信号的贡献率之后, 研究根据贡献率对信号是否保留进行判断。在进行故障检测时, 还可以根据贡献率检测结果进一步确定故障出现位置。如果主成分信号出现异常, 那么可以根据贡献率着重检查对应测点。这样进一步缩减了故障识别方法的运算时间, 提高了风电机设备运行过程故障的检测效率。

1.2 LSTM 序列预测模型的优化策略

在对风电机故障信号的异常状况进行识别之前, 需要对时间序列信号的下一步走向进行预测, 得到理论健康值。若实际测得的信号与预测结果有较大出入, 可初步判断此时该测点信号出现异常。为此, 研究在设备运行过程中利用机器学习方法对传感器检测到的时间序列信号进行预测。目前的时间序列预测模型有很多, 但预测精度都有较大的提升空间。而故障识别的前提就是精确的序列预测, 因此研究对序列预测模型的优化与改进展开探索。长短时记忆网络 (LSTM, long, short-time memory network) 作为循环神经网络衍生一种神经网络, 其具备较为优良的可记忆性功能。LSTM 能够获取时间序列信号中的非线性特征, 常用于序列信号的分析预测当中^[17-19]。LSTM 的网络结构包括遗忘门、输入门、输出门 3 种结构。遗忘门主要判断是否保留前一刻细胞的状态信息。而输入门从输入信息和前一刻时刻状态以及遗忘门输出的信息中学习, 并决定即将输出的值以及当前细胞的候选状态。输出门根据输入值、当前时刻细胞候选状态以及前一刻隐藏层状态计算当前时刻的细胞状态以及输出值。但传统的 LSTM 只能提取序列信号中的前向信息, 信号信息并不能被充分应

用, 为此, 研究对 LSTM 进行改进, 改进后的 LSTM (I-LSTM, improved long short-term memory network) 流程如图 2 所示。

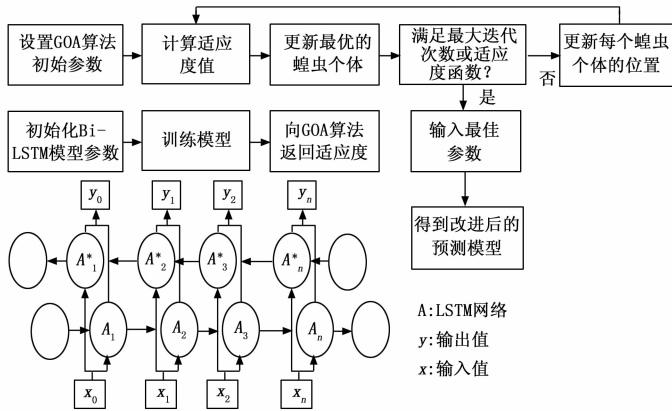


图 2 I-LSTM 的具体流程

为了提高 LSTM 的性能, 研究在 LSTM 基础结构上增加双向结构, 保证模型能够在提取前向信息的同时, 获得后序信号^[20-21]。前向 LSTM 和后向 LSTM 的结构基本相同, 区别在于, 后向 LSTM 是以倒序的方式将信号输入网络中。研究使用加入双向结构的 LSTM 对序列信号进行预测分析, 双向的 LSTM 网络的计算方法如式 (7) 所示:

$$\begin{cases} \vec{h}_t = [\vec{h}_t, \tilde{h}_t]^T \\ \vec{h}_t = F(\omega_t, h_{t-1}) \\ \tilde{h}_t = B(\omega_t, h_{t-1}) \end{cases} \quad (7)$$

式 (7) 中, \vec{h}_t 和 \tilde{h}_t 分别为前向和后向 LSTM 网络在 t 时刻的隐藏层输出, F 和 B 为反方向的两个输入操作, h_t 为 \vec{h}_t 和 \tilde{h}_t 两者的融合输出。参数的选择与模型的预测效果有着直接的影响, 若参数设置合适将会进一步改进模型的预测效果。为进一步提高模型的精度, 优化模型参数, 研究使用蝗虫优化算法 (GOA, grasshopper optimization algorithm) 对模型参数进行寻优。GOA 算法是一种以蝗虫觅食思想为理论指导的启发式优化算法。每个蝗虫个体作为搜索空间的搜索个体, 在空间中进行参数寻优。在 I-LSTM 模型的参数寻优过程中, 研究利用 GOA 算法对模型的权重和偏置等参数进行优化。首先, 初始化蝗虫个体, 每个个体代表一组 LSTM 模型的参数。然后, 根据每个个体在训练集上的预测误差情况来评估其适应度。接着, 通过模拟蝗虫的觅食和聚集行为, 更新个体的位置, 即调整参数值, 并选择出适应度更高的个体。最后将收敛得到的最优参数组合应用到 I-LSTM 模型中。与传统的随机搜索或梯度下降等优化方法相比, 研究选择的 GOA 算法具有更好的全局搜索能力和鲁棒性, 能够更有效地处理复杂且多维的 LSTM 模型参数优化问题。且 GOA 算法还具有良好的并行性和可扩展性。通过结合 GOA 算法对预测模型进行优化, 能够有效提高风电机故障信号序列预测的精度, 使得模型在实际应用中具有更好的泛化能力和鲁棒性。种群位置的变化过程如式 (8) 所示:

$$X_i = S_i + G_i + A_i \quad (8)$$

式 (8) 中, X_i 为种群位置, S_i 为种群个体之间对第 i 个个体的影响, G_i 为第 i 个蝗虫受到的重力影响, A_i 为第 i 只蝗虫受到的风力作用。在初始化种群之后, 需要计算每只蝗虫个体的适应度函数值, 并将最优蝗虫个体寄存在变量中。根据式 (9) 更新参数:

$$c = c_{\max} - t' \frac{c_{\max} - c_{\min}}{T_{\max}} \quad (9)$$

式 (9) 中, t' 为当前的迭代次数, c 为当前需要调整的参数, T_{\max} 为最大迭代次数, c_{\max} 和 c_{\min} 分别为参数的最大和最小取值。基于式 (10) 更新蝗虫个体的位置并计算每只个体的适应值:

$$X_i^d = c \left(\sum_{j=1, j \neq i}^N c \frac{u_{bd} - l_{bd}}{2} s(|x_i^d - x_j^d|) \frac{x_i - x_j}{d_{ij}} \right) + \hat{T}_d \quad (10)$$

式 (10) 中, x_i^d 为当前最优个体, d 为空间维数, 即优化参数的个数, u_{bd} 和 l_{bd} 为第 d 个优化参数的上界和下界, \hat{T}_d 为种群在 d 维空间的最优解。

1.3 基于 I-LSTM 的风电机故障诊断模型

目前建立的异常识别方法普遍对异常情况讨论并不充足, 对异常的判断较为简单, 容易出现误诊漏诊的情况。同时, 工业机械设备的工况复杂多变, 如果仅仅对一个工况进行识别, 缺少随机应变的能力, 识别效果也不佳。为此, 研究在得到的改进数据预测模型后, 为了更好地提高目前故障识别模型的性能, 研究根据 I-LSTM 预测模型得到的预测结果与实际测得信号进行比较, 获得残差。若残差阈值大于设定阈值, 则判定该测点存在故障或即将出现故障。利用预测残差阈值的异常信号识别策略如图 3 所示。

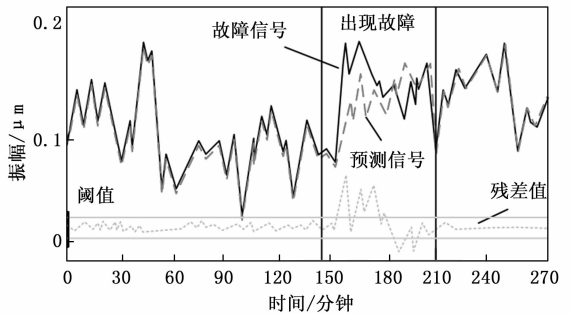


图 3 利用预测残差阈值的异常信号识别策略

目前利用预测残差进行信号异常识别主要使用残差区间以及预警阈值两种。第一种根据残差序列的均值和标准差对残差区间进行划定。若均值和标准差其中一个数值或两个数值都脱离了正常范围, 那么说明信号出现了异常。但这种方法划分的区间比较大, 并不能够及时实现故障预警, 帮助工作人员及时对设备进行维护。第二种方法是对残差阈值进行提前设置, 若过大将预警, 但这种方法并不适用于多种工况及设备类型的情况。为此, 研究将贝叶斯理论引入到异常识别当中, 把所有的未知量都当成随机变量, 探寻其概率分布。该方法能够将先验信息与当前信息

进行结合分析,降低数据中随机性对分析结果的影响。研究利用贝叶斯理论得到数据的先验分布与后验分布,两者分别依靠历史信息 and 当前样本获得。贝叶斯公式将先验和后验分布建立了联系,具体如式(11)所示:

$$\pi(\theta | x) = \frac{h(x, \theta)}{m(x)} = \frac{P(x | \theta)\pi(\theta)}{\int_{\Theta} P(x | \theta)\pi(\theta)d\theta} \quad (11)$$

式(11)中, $\pi(\theta)$ 和 $\pi(\theta | x)$ 分别为先验分布和后验分布, $P(x | \theta)$ 为样本的总体分布, $m(x)$ 为样本 x 的边缘密度函数, θ 为信号参数。为检验所得均值和方差值是否存在异常,研究建立了一个滑动窗口,用以获取对应时间点的预测残差以及历史残差信息。将这些残差信息组合成一个残差集,对信号发生异常概率进行计算。表现为拒绝原假设,接受备选假设的概率即为信号出现异常的概率。在确定均值和标准差的后验概率分布之前需要确定其先验概率分布,两者的分布如式(12)所示:

$$\begin{cases} \sigma^2 \sim IGa(v/2, \nu\rho^2/2) \\ \mu | \sigma^2 \sim N(\tau', \sigma^2/\delta) \end{cases} \quad (12)$$

式(12)中, μ 和 σ^2 分别为均值和方差, ν 为自由度参数, ρ 为尺度参数, τ' 为位置参数, $N(\cdot)$ 为正态分布, $IGa(\cdot)$ 为逆伽马分布, δ 为修正系数。方差的独立后验分布以及其和均值的条件后验分布如式(13)所示:

$$\begin{cases} \sigma_n^2 \sim IGa(v_n/2, \nu_n\rho_n^2/2) \\ \mu | \sigma_n^2 \sim N(\tau'_n, \sigma_n^2/\delta_n) \\ v_n = v + n \\ \nu_n\rho_n^2 = \nu\rho^2 + (n-1)s^2 + \frac{\delta n(\tau' - \bar{e})^2}{\delta + n} \\ \tau'_n = \frac{\delta\tau' + n\bar{e}}{\delta + n} \\ \delta_n = \delta + n \end{cases} \quad (13)$$

式(13)中, \bar{e} 为预测值与真实值之间的平均残差, s 为修正方差。在获得后验概率分布之后,研究根据所得概率分布结果得到信号数据在健康状态下的估计值,对比分析识别是否出现异常。综上,研究首先使用传感器对风电机设备运行数据进行采集,然后利用箱型图法、小波包降噪等对采集数据进行预处理。之后使用 I-LSTM 对信号序列进行预测,并将预测结果与实际测得信号进行比较,得到预测残差。最后根据残差预测结果,并结合贝叶斯理论进行异常信号识别。由此得到故障识别结果,帮助工作人员及时了解风电机设备运行情况,并对设备进行维护。

2 基于 I-LSTM 的风电机设备故障诊断模型性能分析

2.1 模型性能分析

为检验研究设计的设备运行故障诊断模型的性能,研究设计了一系列实验对其进行检验。研究以某风电场中的直驱式风电机为研究对象,其 2023 年 6 月至 2023 年 12 月的故障维修数据以及风速、转速、功率等 42 个特征变量的时间序列数据进行预测分析。研究利用振动传感器、速度传感器、压力传感器以及位移传感器等物理型传感器每 10 分钟记录一次该设备的状态数据。测得的信号包括设备的

振动信号、速度变换信号、设备压力信号、叶片转速信号等。选取实际生产过程中的 6 454 组数据为训练集,4 332 组数据为测试集,正常数据与故障数据的比值为 10:1。研究设备以及实验装置的具体情况如图 4 所示。

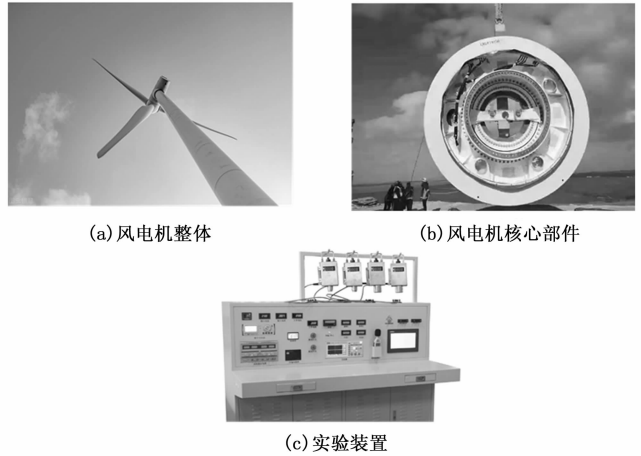


图 4 研究设备及实验装置实物图

为检验研究使用的利用 GOA 算法对 I-LSTM 进行优化的时间序列预测模型的预测准确性,研究选取极限学习机(ELM, extreme learning machine)、深度卷积神经网络(DCNN, deep convolutional neural networks)、季节性差分自回归滑动平均模型(SARIMA, seasonal autoregressive integrated moving average)与研究模型进行对比。4 个模型在相同环境下,对研究数据进行预测,并对比实际数据结果进行对比分析。为公平地评价预测模型的性能,研究使用均方根差(RMSE, root mean square error)对模型的预测误差进行判定。此外,为更加准确量化模型的预测精度,研究利用纳什系数对预测精度进行评定。对模型预测结果进行分析。研究通过估计先验残差集的均值和方差,计算出故障置信度。置信度的计算方法如式(14)所示。

$$\lambda = \pi(\mu | \bar{e}) \quad (14)$$

式(14)中, \bar{e} 为残差平均值, μ 为信号均值, λ 为置信度。具体结果如图 5 所示。

由图 5 (a) 可知,使用研究设计的模型进行预测后得到的数据,在残差大的地方,置信度就较低,相反在置信度较高的地方,残差都较小。这说明研究模型的局部预测精度较为理想。由图 5 (b) 可知,相较于 ELM、DCNN 和 SARIMA 模型,研究构建的模型的拟合效果最好,平均 RMSE 最小,值为 5.142。而 ELM 的平均 RMSE 值为 13.465,相较于研究模型多 8.323; DCNN 的平均 RMSE 值为 10.855,相较于研究预测模型多 5.713; SARIMA 模型的平均 RMS 值为 8.451,其曲线与研究设计的预测模型最为接近,但误差值明显高于研究设计模型。由图 5 (c) 可知,GOA-I-LSTM 的纳什系数数值为 0.985,相较于其他模型明显更高。综合上述内容可知,研究设计的时间序列预测模型能够实现较高精度的序列预测,为后续异常信号识别提供精确数据支持。

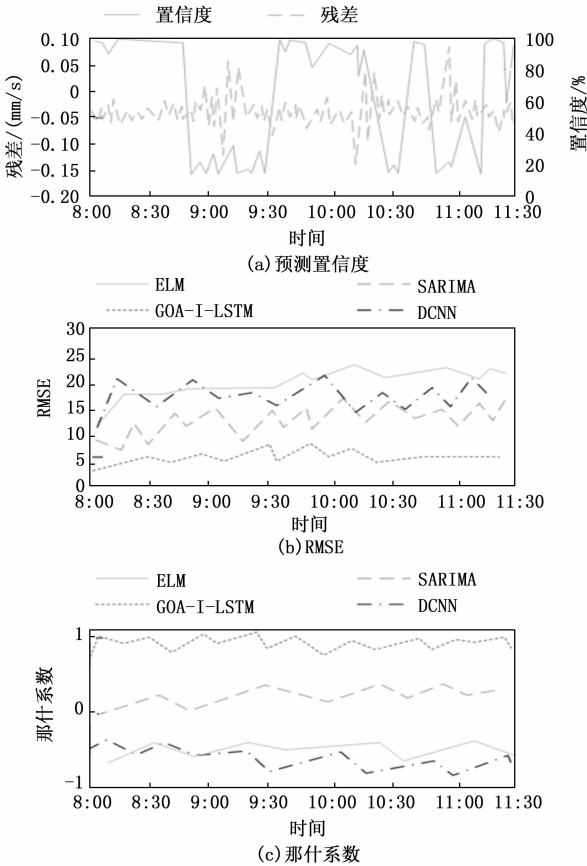


图 5 4 个模型的预测结果对比分析

2.2 模型应用效果实例分析

为进一步检验模型的实际应用效果，研究将模型应用于风电场的日常监控当中。2023 年 11 月，风电场的直驱式风电机组突然发生异常振动，并伴随功率输出不稳定的现象。研究使用本模型对其进行故障诊断。研究将风电机组近段时间各项运行数据进行采集，在经过预处理之后对数据进行预测。模型预测结果显示，风电机组在出现异常前 6 天开始，其功率输出与预测值存在较大的偏差，且预测残差在该时间段内明显增大。这表明风电机组在该时间段内可能出现了异常。为进一步确定异常的类型和位置，研究使用模型对预测残差进行了后验概率分布分析。分析结果显示，异常主要发生在风电机组的齿轮箱和发电机部分。同时，通过对比不同时间段的数据，还发现异常发生的原因与齿轮箱内的轴承磨损和发电机内部的电气故障有关。基于分析结果，工作人员及时对风电机组进行维修。在之后，研究将模型投入到风电机组的使用中。在使用之前，研究对所设计设备策略可靠性进行评估，将投入使用的风电机组前 4 个月的运行数据作为基准，进行预警，并结合实际的维修情况与预警时间及次数的准确性进行评估。结果如图 6 所示。

如图 6 (a) 可知，研究所设计的识别策略在此期间出现了 10 次预警，除预警阶段，其他时间的信号变化情况正常。如图 6 (b) 可知，策略准确预测了 10 次警报，漏诊率

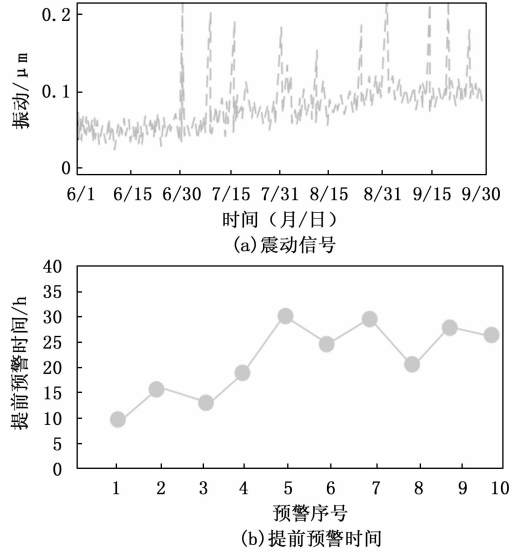


图 6 研究所设计识别策略的实际应用

0%。且提前预警时间基本上在 7 个小时以上，能够较早提醒工作人员对相应机械设备进行维护，避免安全事故等发生。此外，根据风电机的实际维修记录可以看出，研究设计的识别策略进行预警的几次均正确。由此可以看出研究设计的识别策略能够较为可靠地实现机械设备运行故障的准确诊断及预警。

为进一步检验研究利用机器学习和传感器技术构建的风电机设备运行故障检测与诊断模型（模型 1）的性能，研究将其与基于循环神经网络的故障诊断模型（模型 2）、基于图像处理和半监督学习的设备故障诊断模型（模型 3）、基于改进灰狼算法和支持向量机的设备故障诊断模型（模型 4）同步应用于实际生产过程当中。研究利用几种方法对设备运行过程中出现的异常情况进行预警，但工作人员依旧如以往一样对设备进行维护。在投入使用 6 个月之后，根据风电机 4 种模型的预警时间以及之后设备出现明显故障的时间相减，获得提前预警时间。同时根据工作人员在实际维修中对设备出现故障类型以及位置进行的记录，与模型预警诊断的结果进行比较，将每次诊断准确与否进行记录，在最后统一计算出诊断准确率。在投入使用 6 个月之后，统计出工作人员的实际维修次数与几个模型的预警次数进行比较分析，得到漏诊率。具体结果如表 1 所示。

表 1 4 个模型的故障诊断结果对比

指标		诊断准确率/%	漏诊率/%	平均提前预警时间/h	误诊率/%	报警次数
设备 1	模型 1	98.74	0.00	14	1.26	21
	模型 2	92.55	4.76	9	7.45	20
	模型 3	90.84	0.00	9	9.16	21
	模型 4	91.47	14.29	10	8.53	18
设备 2	模型 1	99.01	0.00	17	0.99	18
	模型 2	90.58	0.00	8	9.42	18
	模型 3	88.74	5.56	10	11.26	17
	模型 4	91.73	5.56	11	8.27	17

由表 1 可知, 模型 1 的平均诊断准确率在 98.88%, 无漏诊情况, 误诊率在 1.5% 以下, 在提前 14 小时以上进行预警。能够实现较为可靠的故障诊断预警, 帮助工作人员对现场设备状况进行及时了解。模型 2 的平均诊断准确率为 91.57%, 相较于模型 1 低 7.31%; 模型 3 的平均诊断准确率为 89.79, 相较于模型 1 低 9.09%; 模型 4 的平均诊断准确率为 91.60, 相较于模型 1 低 7.28%。且其他 3 个模型均出现了漏诊的情况。由此可知, 与现有的智能算法相比, 研究设计的故障诊断模型能够实现更为及时与准确的故障诊断, 在故障检测精度上有着显著优势。对比现有模型, 研究的创新在于通过预测残差阈值, 能够更好地捕捉数据的长期趋势和短期波动, 从而提高预测精度。综合上述内容可知, 研究所构建模型能够在风电机设备运行过程中进行及时与准确的故障判断, 辅助工业生产过程中的设备维护与监测。

3 结束语

随着工业智能化的发展, 工业生产过程中的设备也越来越复杂, 若使用常规的基于计算机视觉的故障检测方法, 很多故障并不能被检测出来。为此, 研究利用设备运行信号进行故障判断, 提出了一种基于机器学习和传感器技术的解决方案。通过构建基于 I-LSTM 的时间序列预测模型, 实现了对设备运行状态的准确预测。对比 3 种主流模型发现, 研究构建的模型的拟合效果最好, 平均 RMSE 最小, 值为 5.142。且研究设计的时间序列预测模型的 NSE 值最接近于 1, 这说明研究所设计的预测模型的拟合效果最好。在实际的应用过程中, 策略准确预测了 10 次警报, 漏诊率 0%。且提前预警时间基本上在 7 个小时以上。为进一步验证模型的性能, 研究对比了目前较为流行的故障诊断模型进行对比分析。结果表明, 模型 1 的平均诊断准确率在 98.88%, 相较于其他模型明显更优。综合上述内容可知, 研究所构建模型具有较高的准确性和可靠性, 有助于工业生产过程中设备维护和监测的实时进行。目前研究在数据采集和处理方面, 尚未实现完全自动化, 需要人工干预。未来研究可以尝试引入自动化数据采集和处理技术, 以进一步提高模型的预测精度和效率。

参考文献:

- [1] 刘 翀. 电网设备故障的智能处置策略规则库构建方法 [J]. 微型电脑应用, 2023, 39 (3): 149-152.
- [2] 张逸凡. 基于大数据挖掘的发电设备状态监测与故障诊断系统研究 [J]. 自动化应用, 2023, 64 (2): 48-50.
- [3] 董家祥, 翟纪宇, 马 昕, 等. 知识驱动的机械设备故障诊断 [J]. 计算机科学, 2023, 50 (5): 82-92.
- [4] 何 晖, 代 萌, 李 雪, 等. 基于 DCNN-SVM 的道岔智能故障诊断方法研究 [J]. 铁道学报, 2023, 45 (9): 103-113.
- [5] 时培明, 焦 阳, 陈 卓, 等. 采用分数阶域 MFL-Net 的机械智能故障诊断方法研究 [J]. 动力工程学报, 2023, 43 (10): 1326-1334.
- [6] 潘晓博, 葛鲲鹏, 钱孟浩, 等. 基于改进联合分布适应的轴承智能故障诊断方法 [J]. 机电工程, 2023, 40 (9): 1354-1362.
- [7] 钱 虹, 孙 勃, 郭媛君, 等. 串行分组深度学习运行状态分析与故障预测 [J]. 控制工程, 2023, 30 (5): 936-943.
- [8] 陈富国, 蔡 杰, 李中旗. 基于机器视觉的高压隔离开关设备状态判别与故障诊断技术 [J]. 微型电脑应用, 2022, 38 (2): 191-194.
- [9] 魏 伟, 赵小强, 吴 进. 基于 VMD-ICSO-GRU 的高铁列车车载设备故障率时间序列预测 [J]. 铁道学报, 2023, 45 (6): 58-68.
- [10] 孙洪展. 基于 GRU 神经网络的机械设备故障预测 [J]. 计算机与数字工程, 2023, 51 (8): 1817-1820.
- [11] 倪 强, 李学明, 刘 侃, 等. 基于时序特征模式识别的列车网侧过流故障实时诊断 [J]. 中国电机工程学报, 2022, 42 (11): 3963-3974.
- [12] PHAM Q H, LONG C T, KANIT T, et al. Mechanical performance and failure analysis of a fabric subjected to ballistic impact using multi-scale numerical modelling: [J]. Journal of Composite Materials, 2023, 57 (7): 1331-1342.
- [13] SINGH S, KHARUB M, SINGH J, et al. Brief survey on mechanical failure and preventive mechanism of turbine blades [J]. Materials Today: Proceedings, 2021, 38 (1): 2515-2524.
- [14] LI D, MARSHALL L, LIANG Z, et al. Bayesian LSTM with stochastic variational inference for estimating model uncertainty in process-based hydrological models [J]. Water Resources Research, 2021, 1 (9): 57-59.
- [15] 谢星怡, 张正江, 闫正兵, 等. 基于信号特征提取和卷积神经网络的轴承故障诊断研究 [J]. 计算机测量与控制, 2023, 31 (10): 21-27.
- [16] 王鑫明, 付永波, 李俊涛. 基于决策树算法的电力设备故障诊断辅助决策方法 [J]. 自动化应用, 2023, 64 (12): 128-131.
- [17] 蒋富康, 陆金桂, 刘明昊, 等. 基于 CEEMDAN 和 CNN-LSTM 的滚动轴承故障诊断 [J]. 电子测量技术, 2023, 46 (5): 72-77.
- [18] 陶浩然, 许 昕, 潘宏侠, 等. 基于 EEMD 和 GA-LSTM 算法的行星齿轮故障诊断方法 [J]. 机电工程, 2023, 40 (11): 1700-1708.
- [19] 魏 玮, 吕 游, 齐欣宇, 等. 基于 CNN-LSTM-AM 动态集成模型的电站风机状态预测方法 [J]. 仪器仪表学报, 2023, 44 (4): 19-27.
- [20] 王博乔, 张 彬, 林叶锦, 等. 基于 RAdam I-LSTM 的 LNG 动力船舶上甲板储罐泄漏后果预测方法 [J]. 中国航海, 2023, 46 (2): 60-66.
- [21] 李春生, 田梦晴, 张可佳. 基于 I-LSTM 网络的管道异常数据检测方法 [J]. 计算机技术与发展, 2023, 33 (6): 215-220.