

基于大数据分析的光伏数据监控系统故障诊断应用研究

李晓曼

国网天津市电力公司武清供电分公司 天津 301700

摘要：随着光伏电站规模的持续扩大，传统数据监控与故障诊断方法在实时性、准确性和智能化方面已难以满足运维需求。为提升光伏系统运行的可靠性与发电效率。

关键词：大数据分析；光伏电站；数据监控；故障诊断

引言

随着全球对可持续发展和绿色能源的日益重视，光伏技术在新能源基建项目中的应用已成为研究热点。光伏电站作为清洁能源的重要组成部分，其运行效率和稳定性直接关系到能源转型目标的实现。然而，由于分布式光伏电站数量多、分布广、站点零散等特点，传统监控手段难以满足现代化运维管理的需求，导致故障诊断不及时、发电效率低下等问题频发。在此背景下，大数据分析技术的引入为光伏数据监控系统的故障诊断提供了新的解决方案。通过实时采集、处理和分析光伏系统的运行数据，大数据技术能够显著提升故障识别的准确性和运维决策的科学性。

一、光伏数据监控系统与故障诊断技术概述

（一）光伏数据监控系统架构

光伏数据监控系统作为实现光伏发电系统智能化运维的核心工具，其基本组成主要包括数据采集、数据传输、数据存储与数据分析等环节。这些环节相互协作，共同构成了一个高效且可靠的监控体系。在数据采集环节，通过部署于光伏电站现场的传感器和量测设备，实时获取光伏组件的运行参数，如电压、电流、温度以及环境变量如光照强度和风速等。这些数据为后续的分析提供了基础支持。数据传输环节则负责将采集到的数据通过工业以太网或无线通信技术上传至云端服务器，确保数据的及时性与完整性。在数据存储层面，采用分布式数据库和云存储技术，以满足海量数据的高效管理和长期保存需求。此外，数据分析环节利用大数据技术和人工智能算法对存储的数据进行深度挖掘，生成可视化报告并提供决策支持，从而优化光伏电站的运行效率与维护策略。各环节之间的紧密配合不仅提升了系统的整

体性能，还为故障诊断和预测性维护奠定了坚实基础。

（二）常见故障诊断技术

在光伏数据监控系统中，故障诊断技术的应用至关重要，其发展经历了从传统方法到智能化技术的演进过程。传统的故障诊断技术主要包括基于物理模型的方法和基于信号处理的技术。基于物理模型的故障诊断方法通过建立光伏组件的数学模型来描述其运行特性，并结合实际测量数据进行对比分析，从而定位故障源。然而，由于光伏系统的复杂性，此类方法往往受限于模型的准确性和适用性。基于信号处理的故障诊断技术则通过对光伏电站运行过程中产生的时序数据进行频谱分析，提取故障特征以实现故障检测。例如，在通信故障和设备故障的情况下，发电功率信号的频域特征表现出显著差异，这为故障识别提供了依据。然而，传统方法在处理非线性、高维度数据时存在局限性，难以应对复杂场景下的故障诊断需求。

二、大数据分析在光伏数据监控系统中的应用

（一）大数据采集与预处理

光伏系统的运行数据是进行故障诊断和性能优化的基础，其采集过程涉及多个关键参数的实时监测。传感器网络作为数据采集的核心工具，能够记录光伏板的电压、电流、温度以及光照强度等参数，这些数据以高频时间序列的形式被捕获并上传至监控系统中。在实际应用中，数据采集的频率通常设定为每分钟一次，以确保数据的实时性和准确性。例如，在一项实验中，传感器连续7天采集了约10080条数据，涵盖了不同环境条件下的运行状态，从而验证了系统在各种操作条件下的稳定性和效率。数据采集完成后，需经过一系列预处理步骤以提高数据质量。首先，数据清洗是去除异常值和缺失值的重要环节。由于环境因素或传感器故障可能导致

数据异常，因此需要通过统计方法和领域知识识别并修正这些问题。其次，归一化处理用于消除不同参数量纲之间的差异，使得数据具有可比性。例如，在分析光伏发电功率时，通常采用公式 $P = \frac{p'}{W}$ 对发电功率进行归一化，其中 p' 表示某时刻的实际发电功率， W 表示发电容量，从而得到单位发电容量发电功率的百分比值。这种预处理方法不仅提高了数据分析的准确性，还为后续的特征提取和模型训练奠定了基础。

（二）大数据存储与管理

随着光伏电站规模的扩大和运行时间的延长，其产生的数据量呈指数级增长，这对数据存储和管理提出了更高的要求。分布式数据库和云存储技术因其高可扩展性和可靠性，成为光伏大数据存储的主要解决方案。分布式数据库通过将数据分散存储在多个节点上，能够有效应对大规模数据的存储需求，并提供高效的查询和访问能力。而云存储则利用虚拟化技术实现了资源的动态分配，不仅降低了存储成本，还增强了数据的安全性。

（三）大数据分析应用

在光伏数据监控系统中，大数据分析技术的应用涵盖了从数据挖掘到模式识别的多个层面，为故障诊断提供了强大的技术支持。统计分析是最基础的分析方法之一，通过对历史数据的均值、方差、相关性等统计指标的计算，可以揭示光伏系统运行的基本规律和潜在问题。例如，通过对光伏电站的发电功率曲线进行统计分析，可以发现其与气象条件之间的强相关性，从而为故障预警提供依据。机器学习方法则进一步提升了数据分析的深度和广度。支持向量机（SVM）、随机森林等算法被广泛应用于光伏数据的分类和回归分析中。例如，基于电力公司电表采集的光伏发电数据，可以利用机器学习算法刻画光伏发电功率特征曲线，并通过与基准曲线的对比筛选出异常数据点，进而识别典型故障。此外，深度学习方法在处理复杂非线性关系方面表现出色，尤其是在时序数据分析中展现了显著优势。深度残差网络（ResNet）作为一种先进的深度学习模型，通过引入残差结构解决了梯度消失问题，从而能够更准确地捕捉光伏设备运行时序数据中的细微特征。实验结果表明，基于深度残差网络的故障诊断模型在测试集上的准确率达到98.59%，充分体现了其在光伏故障诊断中的潜力。这些分析方法的综合应用，为光伏数据监控系统的智能化运维提供了坚实的理论和技術支撑。

三、基于大数据分析的光伏数据监控系统故障诊断模型构建

（一）故障特征提取

光伏系统在运行过程中可能面临多种故障类型，包括逆变器故障、通信故障、设备老化等。这些故障类型通常表现为不同的数据特征，例如电压波动、电流异常、功率输出不稳定等。为了准确识别和定位故障，必须首先通过大数据分析提取关键故障特征。文献指出，基于功率信号分析的方法能够有效挖掘光伏电站故障的特征信息，例如通信故障导致的多峰波形或直线波形，以及设备故障引发的接近正弦波的功率曲线。此外，频域分析也揭示了不同故障类型的能量分布特性，如通信故障的能量泄漏到高频段，而设备故障的能量则集中于低频段。通过对历史运行数据的深度挖掘，结合卷积神经网络（CNN）和长短记忆网络（LSTM），可以进一步优化故障特征的提取精度，为后续模型构建提供可靠的基础。

（二）模型选择与训练

在光伏故障诊断领域，多种机器学习模型被广泛应用于故障识别任务，其中包括人工神经网络（ANN）、支持向量机（SVM）以及深度残差网络（ResNet）等。人工神经网络因其强大的非线性映射能力，在处理复杂数据关系时表现出色；支持向量机则以其良好的泛化性能适用于小样本数据集；而深度残差网络通过引入残差结构有效解决了深度模型训练中的梯度消失问题，从而提升了模型的诊断能力。在实际应用中，模型的选择需根据具体场景和数据特点进行权衡。模型训练过程主要包括数据集划分、参数调整以及损失函数优化等步骤。例如，在ResFD-PV模型中，研究者通过设置不同的卷积核大小和残差单元结构，探索了模型性能与计算复杂度之间的平衡点。实验结果表明，当卷积核长度分别为5、5和3时，模型在测试集上的准确率达到98.59%，验证了深度残差网络在光伏故障诊断中的优越性。

（三）模型评估与优化

为了全面评估故障诊断模型的性能，通常采用准确率、召回率以及F1值等指标进行量化分析。准确率反映了模型正确分类样本的能力，召回率衡量了模型对正样本的捕捉能力，而F1值则是综合考虑准确率和召回率的综合评价指标。在模型训练完成后，需通过交叉验证或独立测试集对模型性能进行验证，并根据评估结果进行优化。例如，文献中提到的ResFD-PV模型在训练过程中，通过动态调整学习率、批量大小等超参数，显著提

升了模型的收敛速度和诊断精度。此外,针对模型过拟合问题,可采用正则化技术或增加训练样本数量加以改善。最终优化后的模型不仅具备较高的故障识别准确率,还能在实际应用中保持稳定的性能表现,为光伏系统的运维管理提供有力支持。

四、应用案例分析

(一) 案例背景介绍

本节以某分布式光伏电站项目为例,探讨基于大数据分析的光伏数据监控系统故障诊断的实际应用。该项目位于华北地区,总装机容量为10MW,由多个子站组成,分布于不同的工业园区和建筑屋顶。由于站点分散、环境条件复杂,传统的人工巡检方式难以满足高效运维的需求。为此,项目采用了智能化监控系统,该系统集成了先进的传感器网络、物联网技术和大数据分析平台,能够实时采集光伏组件的运行数据,包括电压、电流、温度及光照强度等关键参数。此外,系统还具备数据存储与管理功能,支持对历史数据的深度挖掘与分析。监控系统的技术架构基于云计算和微服务设计,确保了数据处理的高效性与可扩展性,为后续故障诊断模型的构建提供了坚实的技术基础。

(二) 大数据分析故障诊断过程

在实际应用中,该项目首先通过部署在现场的智能传感器网络采集光伏系统的运行数据,采集频率设定为每分钟一次,以确保数据的实时性和完整性。采集到的原始数据经过预处理后被上传至分布式数据库进行存储,预处理步骤包括数据清洗、缺失值填补以及归一化处理,旨在提高数据质量并减少噪声干扰。随后,利用大数据分析技术对处理后的数据进行深度挖掘,提取与故障相关的特征参数。例如,通过对电压和电流波形的统计分析,识别出异常波动模式;通过机器学习算法构建预测模型,进一步筛选出可能导致故障的关键因素。在故障识别与定位阶段,项目采用了基于深度残差网络的故障诊断模型,该模型能够根据实时数据流快速判断是否存在故障,并精确定位故障发生的具体位置。整个诊断过程实现了从数据采集到故障定位的全自动化操作,显著提升了运维效率。

(三) 应用效果评估

应用基于大数据分析的故障诊断模型后,该光伏项目的运行稳定性得到了显著提升。统计数据显示,系统的平均无故障时间(MTBF)较之前提高了约30%,发电

效率也因故障及时处理而提升了15%以上。此外,智能化运维模式的引入大幅降低了人工巡检频率和维修成本,每年可节省运维费用超过20万元。具体而言,通过预测性维护策略,系统能够在故障发生前发出预警,从而避免因设备损坏导致的停机和生产损失。同时,大数据分析技术的应用还优化了能源管理策略,使得电站的负载分配更加合理,进一步提升了整体经济效益。综合来看,该案例充分证明了基于大数据分析的故障诊断技术在光伏数据监控系统中的实际价值,为行业提供了可借鉴的成功经验。

结论

基于大数据分析的光伏数据监控系统故障诊断应用研究,为光伏电站的智能化运维提供了重要的理论和技术支持。通过构建全面的数据采集、处理与分析框架,本研究实现了对光伏系统运行状态的实时监测与精准诊断。研究表明,大数据驱动的智能监控系统能够显著提升光伏系统的运行效率与可靠性,同时降低运维成本。具体而言,通过对电压、电流、温度等关键参数的实时采集与深度分析,系统能够及时发现潜在问题并预测维护需求,从而有效减少意外停机时间并延长设备使用寿命。此外,基于机器学习和深度学习算法的故障诊断模型在特征提取与模式识别方面表现出色,进一步提高了故障诊断的准确性与时效性。这些研究成果不仅响应了国家“双碳”目标,还为新能源基建项目中的光伏技术优化提供了切实可行的解决方案。

参考文献

- [1] 赵勃. 新能源基建项目中的光伏技术创新与性能优化[J]. 光源与照明, 2024, (2): 141-143.
- [2] 常进. 大数据技术在光伏电站智能监控系统中的应用[J]. 通信电源技术, 2024, 41(11): 243-245.
- [3] 郑伟烁, 郑文悦, 李志伟, 金海薇, 刘雯芳, 马永超. 光伏电站在线监测技术现状与进步趋势展望[J]. 电测与仪表, 2021, 58(9): 1-7.
- [4] 郭凯华, 丁宁, 胡俊葵, 董明. 光伏智能监控平台设计与应用[J]. 设备管理与维修, 2023, (22): 150-151.
- [5] 唐克洪, 周慧, 徐俊. 分布式光伏电站智能化运维系统研究与应用[J]. 科技创新与应用, 2023, 13(19): 193-196.