# 基于电气试验数据的主变压器故障诊断方法研究

李　晓
国网黄冈变电运检分公司　湖北黄冈　438000

**摘　要：**常规的主变压器故障诊断方法以综合特征为主，故障信息存在冗余、错失的问题，影响故障诊断的准确性。本文针对主变压器故障诊断问题，基于电气试验数据展开研究。首先介绍了主变压器故障诊断的重要性和现状，然后详细阐述了电气试验数据的获取与处理方法，进一步探讨了基于不同数据处理技术的故障诊断模型构建，包括特征提取、模式识别与数据分析等关键技术。最后，对研究内容进行了总结，并对未来发展方向提出了展望。本文旨在为基于电气试验数据的主变压器故障诊断提供理论支持和技术参考，以提高主变压器故障诊断的准确性和效率。

**关键词：**主变压器；电气试验数据；故障诊断；特征提取；数据分析

## 引言

主变压器是电力系统中不可或缺的重要设备，其运行状态直接影响到电力系统的安全、稳定和经济运行。随着电力系统的不断发展和设备规模的日益扩大，主变压器故障的诊断与预防成为电力行业关注的焦点。准确、及时地诊断主变压器故障，不仅可以减少停电时间，降低经济损失，还能有效提高电力系统的可靠性。近年来，随着电气试验技术的不断发展，基于电气试验数据的故障诊断方法逐渐成为研究热点。电气试验数据能够反映主变压器在运行过程中的各种状态信息，通过对这些数据的分析和处理，可以实现对主变压器故障的有效诊断。本文将围绕电气试验数据展开研究，探讨主变压器故障诊断的有效方法。

## 1. 电气试验数据的获取

### 1.1 数据来源

主变压器的电气试验数据是故障诊断的基础，其主要来源于定期的预防性试验和在线监测系统。预防性试验是离线检测的重要手段，包括绝缘电阻测试、直流电阻测试、介质损耗因数测试等。这些试验能够深入评估主变压器的绝缘和电气性能，为设备的健康状态提供关键信息。绝缘电阻测试可以检测绝缘材料的老化和受潮情况；直流电阻测试用于评估绕组的连接质量和导电性能；介质损耗因数测试则能反映绝缘介质的损耗特性。与此同时，在线监测系统通过实时采集主变压器运行过程中的电流、电压、油温、油色谱等数据，为故障诊断提供动态信息。

### 1.2 数据采集设备

数据采集设备的性能是确保电气试验数据质量的关键因素。目前，常用的采集设备包括高精度的电压互感器、电流互感器、温度传感器和色谱分析仪等。这些设备能够将主变压器运行过程中的各种物理量转化为电信号，并通过数据采集卡传输到计算机系统中进行存储和分析。电压互感器和电流互感器用于精确测量主变压器的电压和电流，为电气性能分析提供基础数据；温度传感器可以实时监测设备的运行温度，帮助评估热故障风险；色谱分析仪则通过对变压器油中溶解气体的分析，检测绝缘油的老化和故障情况。为了确保采集数据的准确性和可靠性，这些设备需要定期进行校准和维护，以避免因设备老化或故障导致数据偏差，从而影响故障诊断的准确性。

## 2. 故障特征提取方法

### 2.1 基于统计分析的特征提取

统计分析方法是故障诊断中一种基础且重要的特征提取手段。通过对电气试验数据的均值、方差、偏度和峰度等统计量进行计算，能够提取反映主变压器运行状态的关键特征信息。均值反映了数据的平均水平，是数据集中趋势的体现；方差则描述了数据的离散程度，反映了数据的波动情况。偏度和峰度进一步刻画了数据分布的形态，偏度衡量数据分布的对称性，而峰度则反映了数据分布的尖峭或平坦程度。

### 2.2 基于时频分析的特征提取

时频分析方法是一种先进的特征提取技术，能够同时考虑电气试验数据的时间和频率特性，特别适用于处理非平稳

信号。在主变压器的运行过程中，电气试验数据往往具有复杂的时变特性，传统的频域分析方法难以全面捕捉其特征。小波变换作为一种常用的时频分析工具，通过选择合适的小波基函数，可以将电气试验数据分解为不同频率段的信号，并提取各频率段的特征信息。这种分解方式能够有效捕捉信号在不同时间尺度上的细节变化，从而实现对主变压器故障的敏感特征提取。

2.3 基于信息熵的特征提取

信息熵是一种衡量数据不确定性的指标，能够有效反映电气试验数据的复杂程度。在主变压器故障诊断中，通过对电气试验数据的信息熵进行计算，可以提取出与故障相关的特征信息。正常运行状态下，主变压器的电气试验数据通常表现出较低的信息熵，因为其信号较为稳定和规律。然而，当主变压器发生故障时，数据的不确定性会显著增加，信息熵也会随之发生变化。这种变化可以作为故障诊断的有效特征指标。信息熵不仅能够捕捉数据的随机性和复杂性，还能在一定程度上反映故障的严重程度。

## 3. 故障诊断模型构建

3.1 基于机器学习的故障诊断模型

机器学习方法通过建立输入特征与输出标签之间的映射关系，实现对主变压器故障的诊断。在实际应用中，常用的机器学习算法包括支持向量机（SVM）、神经网络和决策树等。支持向量机通过寻找最优分类超平面，将不同故障类型的样本进行分类，尤其在处理小样本、高维数据时表现出色，能够有效避免过拟合问题。神经网络则具有强大的非线性拟合能力，能够自动学习电气试验数据中的复杂关系，适用于复杂的故障模式识别。决策树通过构建树形结构，实现对故障类型的快速判断，其优点在于模型可解释性强，计算复杂度低，适合实时诊断场景。这些机器学习方法各有优势，可根据实际数据特点和诊断需求灵活选择。

3.2 基于深度学习的故障诊断模型

深度学习方法通过构建多层神经网络结构，能够自动提取电气试验数据中的深层次特征，无需人工设计特征提取器。卷积神经网络（CNN）可以对电气试验数据进行卷积操作，提取局部特征，尤其适用于处理图像或高维矩阵数据，能够有效捕捉数据中的空间特征。循环神经网络（RNN）及其变体长短期记忆网络（LSTM）则能够处理时序数据，捕捉电气试验数据的时间序列特征，适用于分析动态变化的故障模式。深度学习模型在大规模数据集上表现出色，具有较高的诊断精度和泛化能力，但需要大量的计算资源和训练数据来优化模型参数，同时对数据的预处理和标注要求较高。

3.3 基于融合模型的故障诊断

融合模型通过将多种机器学习或深度学习方法进行组合，充分利用各方法的优势，提高故障诊断的准确性和鲁棒性。例如，可以将支持向量机与神经网络进行融合，先通过支持向量机进行初步分类，再利用神经网络对分类结果进行优化，从而充分发挥两者的优点，提升诊断性能。此外，还可以将卷积神经网络与长短期记忆网络进行融合，同时提取电气试验数据的空间和时间特征，实现对主变压器故障的精准诊断。这种融合方法不仅能够弥补单一模型的不足，还能在复杂故障场景下提供更可靠的诊断结果，为电力系统的安全运行提供有力保障。

## 4. 数据分析技术在故障诊断中的应用

4.1 数据聚类分析

数据聚类分析是一种无监督学习方法，广泛应用于主变压器故障诊断中。通过将电气试验数据划分为不同的簇，聚类分析能够揭示数据中的内在结构和模式。常用的聚类算法包括 K-Means、层次聚类和 DBSCAN 等。K-Means 算法通过划分簇中心，将数据点分配到最近的簇中，适用于大规模数据集；层次聚类则通过构建树状结构，逐步合并或分割数据点，能够发现数据的层级关系；DBSCAN 算法基于密度进行聚类，能够识别出不同密度的簇，适合处理噪声数据。聚类分析可以将主变压器的运行状态分为正常状态和不同故障状态，为故障诊断提供参考依据。

4.2 数据关联分析

数据关联分析是通过对电气试验数据中不同特征之间的关联关系进行挖掘，以发现隐藏在数据中的故障规律。关联规则挖掘是数据关联分析中常用的方法，通过计算特征之间的支持度和置信度，找出频繁出现的特征组合。支持度表示特征组合在数据中出现的频率，而置信度则反映了特征之间的关联强度。例如，当某两个特征同时出现时，主变压器可能出现某种特定的故障类型。通过关联分析，可以快速识别出与故障相关的关键特征组合，为故障诊断提供重要线索。

4.3 数据可视化分析

数据可视化分析是将电气试验数据以图形或图表的形

式展示出来，直观地呈现数据中的信息和规律。常用的可视化方法包括散点图、折线图、柱状图和热力图等。散点图可以展示特征之间的关系，折线图用于呈现数据随时间的变化趋势，柱状图则用于比较不同类别数据的大小，热力图能够直观地显示特征之间的相关性。通过对电气试验数据的可视化分析，技术人员可以快速发现数据中的异常点和趋势变化，为故障诊断提供直观的依据。

## 5. 故障诊断方法的评估与优化

### 5.1 诊断方法的评估指标

评估故障诊断方法的性能是验证其有效性的重要环节。在主变压器故障诊断中，常用的评估指标包括准确率、召回率、F1 值和 ROC 曲线下面积等。准确率是诊断结果中正确判断的样本比例，反映了诊断方法的精确性；召回率则表示实际故障样本中被正确诊断的比例，体现了诊断方法的完整性。F1 值是准确率和召回率的调和平均值，能够综合反映诊断方法的性能，尤其在数据不平衡时更具参考价值。ROC 曲线下面积（AUC）则表示诊断方法在不同阈值下的性能指标，值越接近 1，诊断方法的性能越好，这一指标广泛应用于故障诊断方法的性能评估中。

### 5.2 诊断方法的优化策略

在变压器故障诊断中，优化策略是提升诊断精度和效率的关键。首先，特征选择优化能够显著提高诊断模型的性能。通过采用 F-score 和信息增益等方法对故障特征进行筛选，可以有效去除冗余特征，保留对故障诊断最具代表性的特征信息，从而提高模型的输入质量。此外，结合改进的特征选择算法，如基于核主成分分析（KPCA）的方法，能够进一步提升特征的代表性，为模型训练提供更高质量的输入。其次，模型参数优化是确保诊断模型性能的重要手段。例如，采用改进的黑猩猩优化算法（ICOA）对最小二乘支持向量机（LSSVM）的惩罚因子和核参数进行优化，能够有效提升模型的精度和泛化能力。此外，通过改进的猎食者优化算法对深度信念网络（DBN）的内部参数进行优化调整，也能显著提高模型的诊断精度。最后，算法改进是优化诊断方法的重要方向。例如，结合多源异构数据融合技术与改进的时间卷积网络（TCN）模型，可以充分利用油色谱分析、红外检测、超声波检测等多种数据源的优势，提升故障诊断的全面性和准确性。

### 5.3 诊断方法的可靠性分析

在变压器故障诊断中，可靠性分析是验证诊断方法在实际应用中稳定性和可信度的重要手段。首先，交叉验证是评估模型性能稳定性的重要方法。通过将变压器试验数据集划分为多个子集，进行多次训练和测试，可以有效评估模型在不同数据划分下的性能。例如，采用 10 折交叉验证方法对改进的 BiLSTM 模型进行验证，结果显示该模型在不同数据集上具有较高的稳定性和可靠性。其次，置信区间估计为诊断结果提供了可信度评估。通过对诊断结果的置信区间进行计算，技术人员可以更好地判断诊断结果的可靠性。例如，基于改进的神经网络方法，通过计算特征重要性及其置信区间，能够确定哪些特征对诊断结果的影响更大，从而为诊断结果提供更可靠的依据。最后，敏感性分析用于评估诊断方法对输入数据的敏感程度。

## 6. 结语

本文围绕基于电气试验数据的主变压器故障诊断方法展开研究，系统探讨了从数据获取到故障诊断模型构建的全过程。通过对数据来源、采集设备及预处理方法的分析，为故障诊断提供了高质量的数据基础。进一步，基于统计分析、时频分析和信息熵的特征提取方法，能够有效挖掘数据中的关键信息，为后续诊断模型的构建提供有力支持。在诊断模型方面，结合机器学习、深度学习及融合模型的优势，提出了多种适应不同场景的故障诊断方法，并通过评估指标、优化策略和可靠性分析，确保了诊断方法的准确性和实用性。

**参考文献：**

[1] 张天祐 . 基于电气试验数据的主变压器故障诊断方法 [J]. 电气技术与经济 ,2024,(10):393-395.

[2] 王钰 , 徐璠 . 综合利用油化和电气试验数据进行变压器故障诊断 [J]. 中国石油和化工标准与质量 ,2024,44(24):28-30.

[3] 张浩 , 盛志华 , 龚伟亮 , 等 . 融媒体环境下基于 SaaS 架构的智能新闻线索平台建设 [J]. 广播与电视技术 .2022 ,49 (07) :45-49.

[4] 郭卫霞 , 张伟 , 杨国玉 . 基于深度级联网络的入侵检测算法研究 [J]. 电子技术应用 . 2021,47 (11) :68-72.

**个人简介：**

李晓 (1991-)，男，湖北浠水，硕士研究生，工程师，研究方向为电力设备检修及试验。