

引用格式: 王童,王正,安丰柱.基于声纹数据标准化的变压器质量缺陷检测研究[J].标准科学,2026(1):71-79.  
WANG Tong,WANG Zheng, AN Fengzhu.Research on Power Transformer Defect Detection Based on Acoustic Fingerprint Data Standardization [J].Standard Science, 2026 ( 1 ) : 71-79.

## 基于声纹数据标准化的变压器质量缺陷检测研究

王童<sup>1\*</sup> 王正<sup>2</sup> 安丰柱<sup>1</sup>

(1.山东国家标准技术审评中心; 2.国网智能科技股份有限公司)

**摘要:**【目的】针对传统声纹检测方法受数据质量不统一及模型泛化能力弱的问题,研究声纹数据标准化方法并构建基于深度学习的质量检测模型,支撑电力变压器无损检测和智能运维。【方法】通过分析变压器声纹特性与缺陷检测瓶颈,构建了涵盖信号采集、降噪处理、特征提取的标准化预处理流程,提升数据质量与一致性。引入基于CNN-Transformer混合架构深度学习模型,实现对多种典型缺陷的识别。【结果】形成涵盖声压级、信噪比、奇偶次谐波比、高频能量占比和谱熵等多维度特征的声纹标准化表征体系,经标准化预处理后的数据能有效提升模型性能,实现对直流偏磁、局部放电等质量缺陷的识别。【结论】本研究为变压器声纹数据提供了标准化处理框架与高精度识别模型,对提升电力设备运维质量具有重要推动意义。

**关键词:** 变压器; 声纹; 数据标准化; 质量缺陷; 检测

DOI编码: 10.3969/j.issn.1674-5698.2026.01.010

## Research on Power Transformer Defect Detection Based on Acoustic Fingerprint Data Standardization

WANG Tong<sup>1\*</sup> WANG Zheng<sup>2</sup> AN Fengzhu<sup>1</sup>

(1.Shandong National Standards Technical Review and Assessment Center; 2.State Grid Intelligence Technology Co., Ltd.)

**Abstract:** [Objective] To address the problem of the inconsistent data quality and weak model generalization, this study aims to investigate acoustic data standardization methods and construct a deep learning-based quality detection model to support non-destructive testing and intelligent maintenance of power transformers. [Methods] By analyzing the characteristics of transformer acoustic signals and the bottlenecks in defect detection, a standardized process covering signal acquisition, noise reduction, and feature extraction is established to improve data quality and consistency. A deep learning model based on a CNN-Transformer hybrid architecture is introduced to identify multiple typical defects. [Results] A standardized acoustic characterization system is established, encompassing multi-dimensional features such as sound pressure level, signal-to-noise ratio, odd-even harmonic ratio, high-frequency energy ratio, and spectral entropy, which can effectively enhance model performance, enabling accurate identification of quality defects such as DC bias and partial discharge. [Conclusion] This research provides a standardized processing framework and a high-precision

**基金项目:** 本文受国家电网公司总部科技项目“电力设备智慧巡检与精准作业机器人关键技术研究”(项目编号: 5108-202218280A-2-249-XG)资助。

**作者简介:** 王童, 通信作者, 博士, 高级工程师, 研究方向为信息与自动化技术标准。

recognition model for transformer acoustic data, contributing significantly to improving the quality of power equipment maintenance.

**Keywords:** power transformer; acoustic fingerprint; data standardization; quality defects; detection

## 0 引言

电力变压器作为电网的核心设备,其运行状态直接关系到整个电力系统的安全、稳定与可靠性。变压器在制造、运输或长期运行过程中,可能因绝缘老化、绕组变形、零部件松动等原因产生各类潜伏性缺陷。这些缺陷在早期通常表现为异常的声音或振动,即声学特征的改变<sup>[1]</sup>。因此,通过对变压器运行时的声纹信号进行分析,实现其质量缺陷的无损、在线检测,已成为状态监测与缺陷检测领域的一个重要研究方向。

电力行业标准DL/T 573—2021《电力变压器检修导则》<sup>[2]</sup>对35kV以上变压器的故障使用拟声词进行了声音特征上的描述,这些描述虽然较为形象但仍然会受到主观判断的干扰<sup>[3]</sup>。得益于计算机和信息技术的发展,机器学习和以深度学习为代表的的人工智能技术在变压器声纹信号缺陷检测方面得到深入研究<sup>[4]</sup>。洪凯星等<sup>[5]</sup>利用支持向量机(Support Vector Machine, SVM)、朴素贝叶斯分类器、二叉决策树等方法对变压器的绕组压紧力衰退状态进行识别。张彬等<sup>[6]</sup>利用小波包变换对不同状态变压器绕组的振动信号进行分析,利用改进后的SVM实现了绕组机械状态的分类诊断。Zhang等<sup>[7]</sup>通过对高低压侧壁的声信号进行采集和筛选,使用反向传播神经网络实现一种变压器内部老化程度估计方法。在深度学习方面,Geng等<sup>[8]</sup>使用语音识别中较为常用的伽马通滤波器倒谱系数构成时频图,并与卷积神经网络相结合实现对10 kV干式变压器机械类故障的诊断。洪凯星<sup>[9]</sup>还对比了SVM、朴素贝叶斯分类器与卷积神经网络在声纹频谱图像识别中的效果,并指出卷积神经网络性能更为优越。包艳艳等<sup>[10]</sup>提出一种基于稀疏表示理论与卷积神经网络(CNN)的检测方法,能够精准识

别750 kV变压器的放电信号。此外,Zollanvari等<sup>[11]</sup>使用循环神经网络架构捕捉声纹时间序列特征,用于判断早期变压器欠励磁、过励磁以及匝间故障的可能性。刘若鹏等<sup>[12]</sup>提出一种基于时延神经网络(TDNN)和多粒度声纹图谱的方法,构建变压器声纹直流偏磁检测模型。

综上所述,传统的变压器声纹检测方法多依赖于专家经验或简单的信号处理技术,结合单一分类模型(如SVM、决策树或浅层神经网络)进行判断,其在实际场景的规模应用仍面临以下挑战:

(1) 声纹数据的复杂性和不一致性。变压器声纹信号受电压等级、设备型号的影响存在显著差异,即便是相同设备在不同负载下其振动和发声强度也不同,同时,噪声、风声、雨声及其他设备运行声会与变压器本体声音混杂,淹没微弱的缺陷特征。

(2) 单一检测模型的局限性。变压器缺陷声纹的特征与类别之间依然存在着复杂的、非线性的映射关系。在面对高维、多变的声纹数据时,依赖单一模型的风险很高,其决策边界可能不够精确和稳定,导致对某些罕见缺陷或模糊样本的误判。

为解决变压器声纹检测中数据质量不一和模型泛化能力弱的难题,本文提出一种基于声纹数据标准化的变压器质量缺陷检测方法。首先研究声纹数据的标准化预处理流程,对音频信号及其归一化、降噪、特征计算等提出要求。其次,采用深度学习算法构建高性能的质量缺陷检测模型。实验结果表明,本文提出的数据标准化流程能够规范数据的特征以适应不同场景,构建的质量缺陷检测模型能够精准检测多类缺陷,为电力设备生产过程中的质量检测、智能运维缺陷识别提供一种更为可靠的技术手段。

## 1 声纹数据标准化预处理

### 1.1 总体流程

在声纹采集过程中,原始信号在采样频率、存储格式以及数据时长方面存在明显差异,同时信号中也可能存在大量无关噪声和干扰,使得后续的声纹分析和质量缺陷检测效果较差。对此,本文提出了声纹的标准化预处理流程,并分别对声纹信号、降噪处理以及特征计算提出要求,旨在提升数据质量与一致性。

图1为声纹数据标准化流程图。高质量的声纹信号是变压器缺陷检测的重要前提。基频与谐波提取基于傅里叶变换将时域信号转换为频域信号,以获取频率域的特征表征。降噪处理的核心目标是在保留变压器本体真实声学特征的基础上,最大限度抑制无关干扰信号。本体声纹信噪比主要用于评估降噪后信号的质量,判断其是否满足特征提取和缺陷检测的要求。奇偶次谐波比等特征与变压器的故障缺陷机理密切相关,蕴含丰富的状态信息。质量缺陷检测算法及实验验证分别在第2章与第3章中详细说明。

### 1.2 声纹基本要求

变压器声纹为音频信号,其编码、存储及采样等不同实现方式直接影响后续的特征分析和缺陷检测。LPCM(线性脉冲编码调制)和PCM(脉冲编码调制)是标准的、非压缩的无损音频编码格式,保留了音频信号的所有信息,为后续的声纹分析提供最纯净、最可靠的原始数据。此外,变压器声纹主要能量集中在100 Hz~600 Hz的低频段,但许多机械性故障(如松动、放电)会产生丰富的高频振动和声音信号。根据奈奎斯特采样定理,较高的采样频率(如32 kHz采样频率最高可分析信号16kHz的频率成分)能够捕获那些对缺陷极其敏感的高频瞬态事件和宽频带噪声。结合实际应用需求,对以下内容进行规范:

- (1) 声纹样本数据存储格式宜选用 LPCM 或 PCM 编码,应以WAV格式进行存储;
- (2) 声纹样本的采样率应不小于32 kHz;

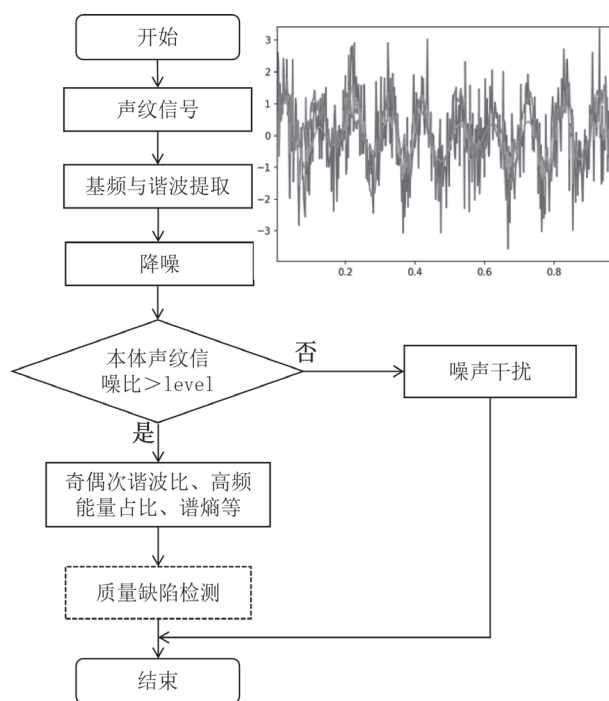


图1 声纹数据标准化流程图

(3) 测试样本集应在实际运行场景采集;

(4) 样本标注信息应至少包括: 变压器型号、标注日期、分类名称、标记开始时间、标记结束时间、样本获取方式、场景及分类名称等, 场景及样本分类;

(5) 将长时间采集的连续音频信号分割为多个固定时长(一般为10 s)的音频片段, 每个片段作为一个独立的样本, 从而扩充样本数量, 并保证每个样本都包含一个稳定的声学事件周期;

(6) 为消除因麦克风增益、距离远近不同导致的声压级绝对值的差异, 对每个音频样本进行幅度归一化, 将其幅度范围缩放至[-1, 1]。其计算公式为:

$$x_{norm}(t) = \frac{x(t)}{\max(|x(t)|)} \quad (1)$$

式中:  $x(t)$  为样本信号;  $\max(|x(t)|)$  为信号的峰值振幅;  $x_{norm}(t)$  表示归一化后的信号。此步骤确保了所有样本处于相同的幅度尺度上, 避免了缺陷检测模型对音量大小的无关关注。

### 1.3 基频与谐波提取

通过快速傅里叶变换(FFT)将时域信号转换为频域信号,便于进行频域特征提取。考虑到变压器声纹基频为100 Hz,且整体为平稳信号,进行FFT计算的数据长度应至少大于1 s,建议数据长度介于1~10 s之间,保证合理的频域分辨率,同时避免冗余计算。

$$X(k) = \sum_{n=0}^{N-1} e^{-i\frac{2\pi}{N}kn} a_n \quad (2)$$

式中:  $a_n$  为第  $n$  个采样点处(或者说在时间  $n$ ) 信号的幅值;  $X(k)$  为在频率处的复数谱;  $N$  为时域总采样点数(大于等于1 s的采样点数)。对某时域信号进行频谱计算效果见图2。

### 1.4 降噪处理的要求

常见的降噪算法包括小波阈值降噪(Wavelet Threshold Denoising)、谱减法、无限脉冲响应(Infinite Impulse Response, IIR)滤波器等。例如,对于干扰频段固定,且与变压器主频及谐波重叠不大的干扰信号,可采用IIR构成窄带滤波器组进行降噪;对于噪声和变压器声纹存在频带重叠,且干扰信号频率分布、幅值特性较为稳定的干扰,可采取谱减法进行噪声抑制。

本文不对具体的降噪算法进行研究,但使用变压器本体声纹信噪比对降噪后的信号进行评价。

### 1.5 本体声纹信噪比

变压器本体声纹信噪比主要用于评价噪声条件下变压器本体声纹的有效性,避免强干扰信号进入后续的质量缺陷检测过程。根据数据标准化预处理流程(见图1),本文的level取值为0.8。当数据噪声较大,并且变压器本体声纹信噪比低于该值时,可能为传感器故障、传感器安装异常、数据采集质量低、强干扰等情形,此时不宜进入缺陷检测流程,直接输出声纹信号为干扰信号。

在基频与谐波提取(公式2)的基础上,依次计算信号声压级(20 Hz~20 kHz)、变压器本体声纹信噪比(50 Hz谐波成分占总声压级比重)。

#### (1) 声压级

$$SPL = 20 \log_{10} RMS / base \quad (3)$$

式中:  $RMS$  为声纹信号有效值(20 Hz以上频段);  $base$  为声压级参考基准,具体为20  $\mu Pa$ 。

#### (2) 本体声纹信噪比

$$SNR = \sqrt{\sum_{i=1}^{60} (\frac{X_{i*50}}{\sqrt{2}})^2} / RMS \quad (4)$$

式中:  $X_{i*50}$  为声纹信号频谱,对应索引为第  $i$  个

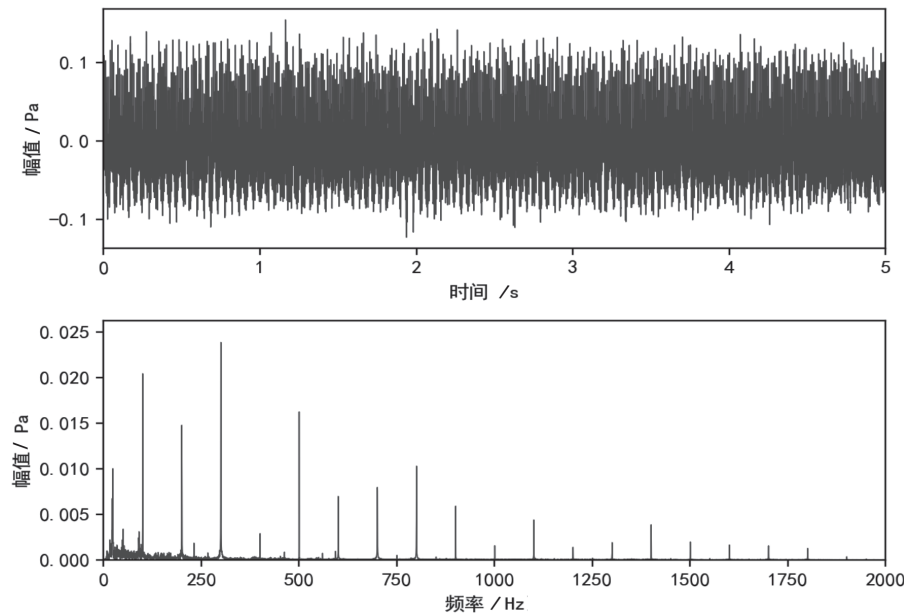


图2 时域图和频谱图(幅值谱)



50 Hz频率。

### 1.6 标准化特征提取

经过对变压器发声机理的研究<sup>[13]</sup>, 奇偶谐波比(50 Hz~2000 Hz内50 Hz奇数倍频和偶数倍频比重)、高频能量占比(800 Hz~2000 Hz能量占50 Hz~2000 Hz频带比重)、共振解调工频相关性(10000 Hz以上频带包络谱中50 Hz和100 Hz成分在50 Hz~500 Hz频带内比重)、谱熵(50 Hz~2000 Hz内工频谐波成分分布熵)等特征与变压器状态密切相关, 可作为变压器缺陷检测的标准化特征。同时信号原始时频特征及Mel时频图可作为深度学习模型的标准输入, 具体各特征的计算过程如下。

#### (1) 奇偶谐波比

$$HR = \sum_{i=1}^{20} X_{(2*i-1)*50}^2 / \sum_{i=1}^{20} X_{(2*i)*50}^2 \quad (5)$$

#### (2) 高频能量占比

$$HFER = \sum_{i=16}^{40} X_{i*50}^2 / \sum_{i=1}^{40} X_{i*50}^2 \quad (6)$$

#### (3) 共振解调工频相关性

$$PD\_CORR = \sqrt{\sum_{i=1}^{10} (\frac{X_{i*50}^e}{\sqrt{2}})^2} / \sqrt{\sum_{i=50}^{500} (\frac{X_i^e}{\sqrt{2}})^2} \quad (7)$$

式中:  $X_i^e$  为声纹信号高频段(通常10000 Hz以上)包络频谱, 索引为对应频率, 该特征主要是通过对变压器声纹信号高频段进行包络解调, 避开低频段工频高次谐波干扰, 然后对包络信号进行FFT计算。

#### (4) 谱熵

$$ENTRONY = \sum_{i=1}^{40} P_i \log_2 P_i \quad (8)$$

式中:  $P_i = X_{i*50} / \sum_{i=1}^{40} X_{i*50}^2$ , 为对应频率处能量占比。

#### (5) 时频图

对音频信号进行短时傅里叶(STFT)变换:

$$X_{t,w} = \sum_{i=0}^{N-1} x_i w_{i-t} e^{-jwi} \quad (9)$$

式中:  $X_{t,w}$  为时间窗口 $t$ 和频率 $w$ 上的STFT值;

$x_i$ 是输入信号的样本值,  $w_{i-t}$ 是窗口函数, 通常使用汉宁窗或矩形窗;  $N$ 是窗口的长度, 根据本场景需要;  $N$ 建议设置为100 ms对应采样点数, 帧移设置为帧长的一半。

#### (6) Mel时频图

$$H_m(f) = \begin{cases} 0, f < x(m-1) \\ \frac{f-x(m-1)}{x(m)-x(m-1)}, x(m-1) \leq f \leq x(m) \\ \frac{x(m+1)-f}{x(m+1)-x(m)}, x(m) \leq f \leq x(m+1) \\ 0, f > x(m+1) \end{cases} \quad (10)$$

式中:  $m$ 为滤波器组编号;  $x(m)$ 为三角滤波器的中心频率:  $x(m) = \left(\frac{N}{f_s}\right) \text{Mel}^{-1}(\text{Mel}(f_{min}) + m \frac{\text{Mel}(f_{max}) - \text{Mel}(f_{min})}{M+1})$ , 其中,  $f_{min}$ 、 $f_{max}$ 分别为滤波器的频率最大值和最小值,  $f_s$ 为声传感器的采样频率,  $N$ 为离散傅里叶变换的帧长度。将上述Mel滤波器组矩阵与时频谱矩阵相乘, 再取对数, 即可得到Mel时频谱。

## 2 基于CNN-Transformer的声纹缺陷检测模型

### 2.1 总体架构

为了有效捕捉变压器声纹时频图中蕴含的缺陷特征, 本文设计了一种结合卷积神经网络(CNN)与Transformer编码器的混合深度学习模型。该模型充分利用了CNN在局部特征提取上的优势以及Transformer在长程依赖关系建模上的强大能力。整体网络结构如图3所示, 主要由CNN特征提取器、位置编码器、Transformer编码器和分类头4个核心模块构成。

### 2.2 输入预处理与数据表示

原始的一维声纹音频信号首先经过短时傅里叶变换预处理模块, 将其转换为二维时频图(公式9)。具体而言, 对信号进行分帧、加汉宁窗处理, 计算其幅度谱, 并转换为分贝尺度(公式10), 以模拟人耳对声音强度的感知。随后, 对频谱进行特定频带的抽取与重排, 最终得到固定尺寸为50×100的时

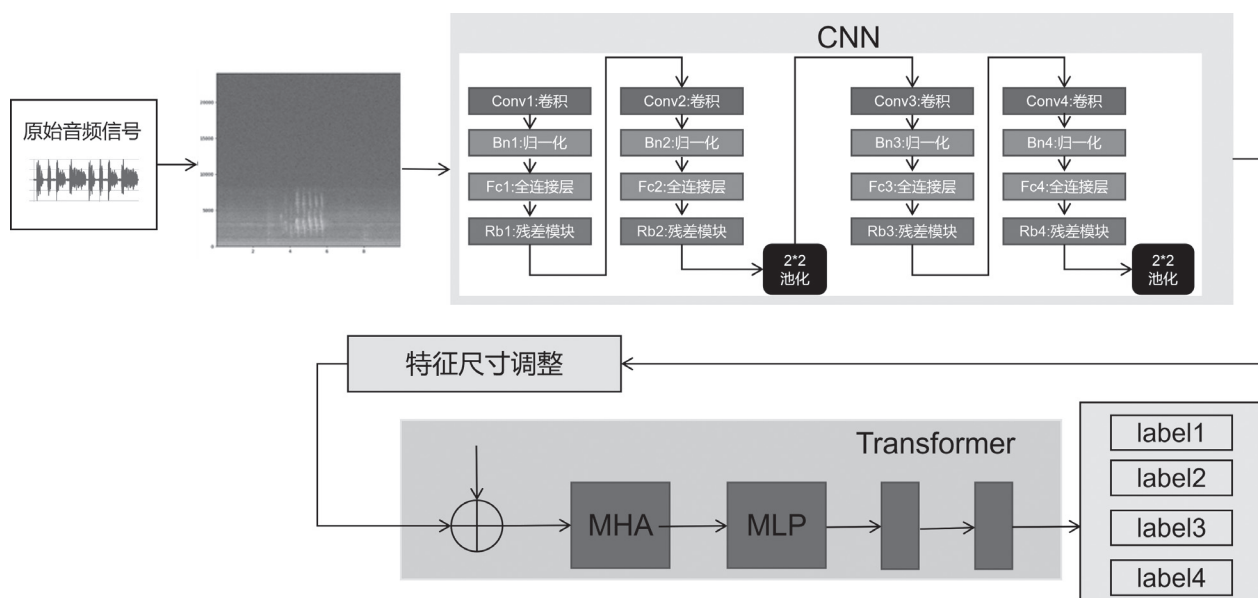


图3 CNN-Transformer架构网络结构示意图

频矩阵。该矩阵以图像的形式表示声音信号，其行代表频率，列代表时间，矩阵值代表特定时刻、特定频率分量的能量强度，以此作为网络的输入。

以信号采样率为32 000 Hz，数据时长为10 s的数据进行Mel时频图计算为例，进行具体说明，MFCC系数数量设置为50，帧长设置为0.2 s，帧移0.1s，该过程将 $1 \times 320000$ 的音频时序数据转换为 $1 \times 50 \times 100$ 的二维时频图像数据，对应图像通道数为1，宽度为100，高度为50。

### 2.3 卷积神经网络特征提取器

CNN特征提取模块负责从时频图中分层地提取由低级到高级的抽象空间特征。该模块采用深度残差结构，以缓解网络加深带来的梯度消失问题，确保训练的稳定性。

输入的单通道时频图（尺寸： $1 \times 50 \times 100$ ）首先经过4个连续的卷积残差块进行处理。每个卷积残差块包含标准卷积层、批归一化（Batch Normalization）层、ReLU激活函数以及一个自定义的残差块（Residual Block）。其中，残差块内部采用两个 $3 \times 3$ 卷积层与一个 $1 \times 1$ 卷积捷径连接的结构，确保输入输出维度匹配。前2个卷积块后不进行下采样，以保留丰富的细节信息；在第二个和第四个卷积块后，插入一个 $2 \times 2$ 最大池化层进行空间下采样，在

提升特征抽象程度的同时有效控制计算复杂度。最终，该模块输出一个 $192 \times 12 \times 25$ 的高维特征图。

此后，特征图被展平并送入一个全连接层，进行特征压缩与重映射。该全连接层将特征维度压缩为128维，同时将展平后的特征序列重新塑形为一个长度为64的序列。这表示CNN模块将二维时频图空间结构转换为了一个一维的特征序列（尺寸： $64 \times 128$ ），其中序列的每一个位置（共64个）都对应一个128维的特征向量，压缩了原始输入的一个特定局部区域的抽象信息。此序列将作为后续Transformer编码器的输入。

### 2.4 位置编码器

由于Transformer本身不具备感知序列顺序的能力，而声纹特征的时序关系至关重要，因此需要在特征序列输入Transformer之前注入位置信息。本文采用正弦和余弦函数生成固定位置编码（Positional Encoding）。该编码与特征序列相加，为序列中的每一个128维特征向量赋予了其绝对位置信息，从而使模型能够充分利用特征在序列中的顺序关系。

### 2.5 Transformer编码器

注入位置信息的特征序列被送入Transformer编码器模块，以建模序列中各个特征向量之间的全局依赖关系。该编码器由N个相同的编码器层堆叠而成

( $N$ 为可配置超参数)。每一层都包含一个多头自注意力 (Multi-Head Self-Attention, MHA) 机制和一个前馈神经网络 (Feed-Forward Network)。

自注意力机制允许序列中的任意一个特征向量与序列中的所有其他向量 (包括自身) 进行交互并计算注意力权重, 从而动态地聚合全局上下文信息。多头机制则将注意力分散到不同的表示子空间, 使模型能够共同关注来自不同位置的不同类型的依赖关系。每个子模块都采用残差连接和层归一化 (Layer Normalization) 来稳定训练过程。

通过多层堆叠, Transformer编码器对输入特征序列进行深层变换, 输出一个同等长度的增强特征序列 (尺寸:  $64 \times 128$ ), 其中每个特征向量都包含了丰富的全局上下文信息。

## 2.6 分类头

为将Transformer输出的序列信息转化为最终的检测分类决策, 首先对整个序列在时间步维度 (长度为64) 上执行全局平均池化 (Global Average Pooling), 将所有时间步的特征信息聚合为一个单一的、128维的全局特征向量。随后, 该向量通过一个线性全连接层, 其输出神经元数目等于缺陷类别的数量 (含正常状态, 本文共6类)。最后, Softmax函数被应用于该层的输出, 生成对应于各个类别的概率分布, 概率最高的类别即为模型的预测结果。

# 3 结果

## 3.1 数据准备

为对本文提出的方法进行验证, 使用5 440条样本组成训练数据集。其中, 4 000条为正常样本, 1 440条为带有缺陷的异常样本; 测试数据集由875条异常样本组成。缺陷类型包括局部放电、直流偏磁、夹件松动、重过载、冷却器异响5类。上述数据样本由电力科研机构通过变压器声纹在线监测和人工巡检方式在实际场景中获取, 面向用户提供算法训练和测试服务。

以上数据经过本文提出的数据标准化预处

理, 其缺陷类别分布见表1。

表1 测试集、训练集中主要缺陷类别分布

类别	正常	局部放电	直流偏磁	夹件松动	重过载	冷却器异响
训练集	4 000	340	240	240	310	310
测试集	/	215	150	150	180	180

## 3.2 评价指标

本文使用准确率对算法进行评价:

$$ACC = \frac{n}{N} \quad (11)$$

式中:  $n$ 为正确识别出的缺陷样本的个数;  $N$ 表示测试集中对应缺陷样本的总数量。

## 3.3 对比算法

作为对比, 本文分别使用特征阈值法、卷积神经网络和本文提出的缺陷检测方法进行测试。特征阈值法是在标准化特征提取的基础上, 与预先设置的阈值参数做比较, 当计算得到的特征值大于阈值时, 则判定发生该类型缺陷。特征阈值法的主要特征、典型缺陷表现、适用缺陷类型及默认参数设置见表2。

卷积神经网络的结构见本文2.3的描述, 未使用位置编码器和Transformer编码器。

## 3.4 实验结果

完成基于CNN-Transformer的声纹缺陷检测模型构建后, 根据实际显存空间设置批处理大小  $batch\_size$  为32, 设置学习率  $lr$  为0.001, 同时可以根据测试损失值动态调整学习率大小, 保证收敛到最佳水平, 设置训练代数  $epochs$  为50。开始第一次训练后, 根据测试集数据表现, 进行网络结构、 $batch\_size$ 、 $epochs$ 、 $lr$  等参数调整, 优化模型参数, 避免模型的梯度消失、爆炸、过拟合、欠拟合等问题, 对模型进行持续深度优化和微调。模型训练过程见图4, 其训练损失和测试损失不断下降, 在  $epoch$  接近30时趋于稳定。

本文在数据标准化预处理的基础上, 对3种方法进行了对比测试。表3是不同方法准确率的实验结果, 准确率计算方法见公式 (11)。其中本文提出的方法和卷积神经网络方法均属于深度学

表2 主要特征含义、用途及默认值

数值特征类型	典型缺陷表现	用于诊断的缺陷类型	阈值参数
声压级（SPL）	重过载缺陷一般表现为声压级显著增大6dB以上	用于重过载缺陷判断	75dB
奇偶谐波比（HR）	出现直流偏磁缺陷时，50Hz奇数倍频显著增大	用于直流偏磁或谐波电流判断	0.4
高频能量占比（HFER）	出现夹件松动缺陷时，高频端信号幅值明显增大	用于判断松动类缺陷	0.4
谱熵（ENTRONY）	出现冷却器异响缺陷时，底噪增大， 频谱复杂度增大	用于冷却器异响类判断	2.8
共振解调工频相关性（PD_CORR）	出现局部放电缺陷时，声纹高频段会存在变压器 工频相关的调制成分	用于局部放电类判断	0.6

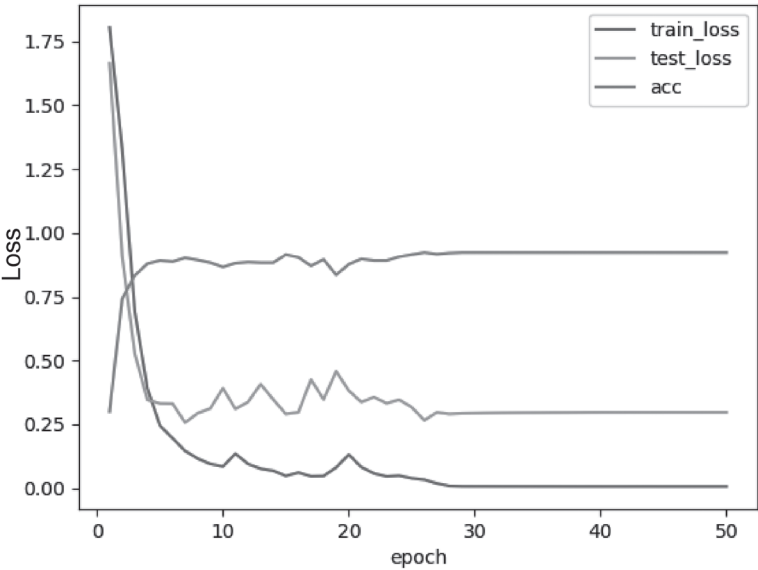


图4 模型训练损失曲线

习检测模型，其检测准确率相近，但由于引入的Transformer网络自注意力机制能捕捉特征序列中任意2个元素之间的依赖关系，有助于理解声音的上下文信息，因此其检测准确率略高于卷积神经网络的方法。总体而言，基于深度学习的检测方法能够挖掘数据之间非线性、高维度特征，有效降低噪声干扰与设备工况变化的影响，在不同类别的缺陷检测上均有较高的准确率，表现出较

强的鲁棒性。  
相比较，特征阈值的方法总体上低于深度学习的方法，对夹件松动和重过载的检测准确率明显低于深度学习的方法，表明该方法对高频特征不敏感，易受到负荷变化、环境噪声的影响，鲁棒性较差。此外，表3中特征阈值方法对直流偏磁的检测准确率最高，表明在样本数据不足的情况下，与发声机理密切相关的特征阈值法在实际应用中仍

表3 准确率结果

方法	局部放电/%	直流偏磁/%	夹件松动/%	重过载/%	冷却器异响/%
特征阈值法	98.14	99.33	92.67	90	97.22
卷积神经网络	99.07	98	96.67	98.33	97.78
本文方法	100	98.67	98.67	100	98.89



具有不可忽视的作用。

## 4 结论

本文针对变压器声纹检测中数据质量不一和检测能力单一的问题,开展了基于数据标准化和深度学习的质量缺陷检测方法研究。主要工作及结论如下:

(1) 提出了1套较为系统的声纹数据标准化预处理流程。该流程规范了声纹数据的采集格式、降噪处理、特征计算与评价,使用本体声纹信噪比作为数据质量评价的指标,有效抑制了强干扰信号对后续检测环节的污染,为构建高质量、高一致性的声纹数据集奠定了基础。

(2) 构建了基于CNN-Transformer深度学习

模型用于变压器缺陷检测。该模型充分发挥了CNN在提取声纹时频图局部空间特征方面的优势,并利用Transformer架构强大的长程依赖建模能力对特征序列进行深层融合与上下文信息提取。实验证明,该混合模型的综合性能优于纯CNN模型和传统的特征阈值法,特别是在局部放电和重过载等缺陷的识别上表现出更高的准确率与鲁棒性。

本文研究结果表明,数据标准化预处理与先进的深度学习模型相结合,能够更好地挖掘声纹数据中蕴含的缺陷特征,是实现变压器状态精准感知与缺陷智能检测的有效途径。未来将继续对声纹数据集的标准化建设开展研究,支撑算法在复杂工况下的泛化能力。

## 参考文献

- [1] 齐子豪,全杰,张中浩,等. 基于多粒度知识特征和Transformer网络的电力变压器故障声纹辨识方法[J]. 中国电机工程学报,2025,45(4):1311-1322.
- [2] 电力行业电力变压器标准化技术委员会. 电力变压器检修导则: DL/T 573—2021[S].
- [3] 李章维,周浩,郑文皓,等. 基于声纹识别的变压器故障检测方法[J]. 浙江工业大学学报,2025,53(3):293-298.
- [4] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2012, 42(2):513-529.
- [5] HONG K X, WANG L, XU S. A variational mode decomposition approach for degradation assessment of power transformer windings[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 68(4):1221-1229.
- [6] 张彬,徐建源,陈江波,等. 基于电力变压器振动信息的绕组形变诊断方法[J]. 高电压技术, 2015, 41(7):2341-2349.
- [7] ZHANG Y, LI J, LIU H, et al. BP-Neural-Network-Based Aging Degree Estimation of Power Transformer Using Acoustic Signal[C]//2020 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD). IEEE, 2020: 617-621.
- [8] GENG Q S, WANG F H, ZHOU D X. Mechanical fault diagnosis of power transformer by GFCC time-frequency map of acoustic signal and convolutional neural network[C]//2019 IEEE Sustainable Power and Energy Conference (iSPEC). IEEE, 2019: 2106-2110.
- [9] HONG K, JIN M, HUANG H. Transformer winding fault diagnosis using vibration image and deep learning[J]. IEEE Transactions on Power Delivery, 2020, 36(2): 676-685.
- [10] 包艳艳,杨广泽,陈伟,等. 基于SBSS与CNN的750 kV变压器和尖板的放电信号声纹识别[J]. 西南交通大学学报, 2025, 60(3):781-792.
- [11] ZOLLANVARI A, KUNANBAYEV K, BITAGHSIR S A, et al. Transformer fault prognosis using deep recurrent neural network over vibration signals[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [12] 刘若鹏,徐辉,赵庆杰,等. 基于多粒度声纹图谱和时延神经网络的变压器直流偏磁故障程度识别方法[J]. 智慧电力, 2025, 53(3):99-106.
- [13] 王博闻. 电力变压器典型异常状态声纹诊断技术[D]. 北京:华北电力大学, 2022.