

基于二维特征提取方法与混合神经网络的接触式采集 110 kV 三相三绕组变压器无载调压异常放电声纹的识别方法

童 昉, 黄文礼*, 李 磊, 晏雨晴

(安徽南瑞继远电网技术有限公司, 安徽 合肥 230088)

A Recognition Method Based on Two-Dimensional Voiceprint Feature Extraction Method and Hybrid Neural Network for Contact-Collected Abnormal Discharge Voiceprint of 110 kV Three-Phase Three-Winding Power Transformer in No-Load Voltage Regulation

TONG Yang, HUANG Wenli*, LI Lei, YAN Yuqing

(Anhui NARI Jiyuan Power Grid Technology Co., Ltd., Hefei 230088, China)

Abstract: Abnormal discharge is a dangerous power transformer fault, which can lead to serious safety hazards if not detected in time. A method for identifying abnormal discharge in transformer is proposed by collecting the voiceprint signal in the transformer box through the contact voice pickup. And a feature extraction method and a deep neural network structure are proposed to achieve efficient identification of abnormal transformer discharge. Firstly, a two-dimensional voiceprint feature extraction method combining Mel frequency and key frequency is designed. Then, a hybrid two-dimensional feature recognition model based on convolutional neural network and Transformer network is used to accurately identify the abnormal discharge voiceprint signal while ensuring the speed. Finally, according to the experimental analysis of the discharge data collected from the 110 kV three-phase three-winding transformer in the no-load voltage regulation process, the recognition speed of the proposed method increases by 0.19 seconds per sample, and the accuracy increases by 4.5% compared with ResNet50.

Key words: abnormal discharge of transformer; voiceprint recognition; voiceprint feature extraction; hybrid neural network

摘 要: 异常放电是电力变压器中一种潜在的危险故障,

基金项目: 新型电力系统智能运维安徽省联合共建学科重点实验室成果

Results of the Anhui Provincial Key Laboratory of Joint Co-construction of New Power System Intelligent Operation and Maintenance

若未及时检测可能导致严重的安全事故。采用接触式拾音器收集变压器箱体内异常放电声纹信号,并提出了一种特征提取方法和一个深度神经网络结构,以实现变压器异常放电的高效识别。首先,设计了一种结合梅尔频率提取和关键频率提取的二维声纹特征提取方法。其次,提出了一种基于卷积神经网络和 Transformer 网络的混合二维特征识别模型,能够在确保识别速度的同时准确辨识异常放电声纹信号。通过对 110 kV 三相三绕组变压器无载调压试验过程中采集的放电数据进行试验分析,所提方法相较于 ResNet50 识别速度增加约 0.16 秒/样本,同时识别效果提升了 4.5%。

关键词: 变压器异常放电; 声纹识别; 声纹特征提取; 混合神经网络

0 引言

电力变压器在输电、变电和配电过程中扮演着至关重要的角色,确保其正常运行是保障电力传输与供给安全的关键环节^[1-4]。随着变压器运行时间的增长,由于运行负荷、设备老化以及绝缘油浑浊等因素的影响,变压器的工况不可避免地受到影响,可能出现铁心松动、绕组变形、匝间短路、直流偏磁和异常放电等故障^[2,4-5]。其中,异常放电是变压器常见故障中危险系数较高的情况之一。多种异常因素,如过负荷、绝缘油污染、绕组故障和散热器故障等,都可能导致异常放电的

发生^[5-7]。异常放电初期局部持续发生,随着放电次数的增多,产生大量热能,导致介质温度升高,甚至局部碳化,可能引发发热击穿风险,导致变压器损坏等严重后果^[6-7]。因此,及时准确地检测异常放电对于变压器的安全运行和整个电力系统的正常运作至关重要。

传统的变压器异常放电监测主要依赖于油色谱的变化观察,尽管该技术准确性较高,但微小的异常放电很难产生足够观测到的化学性质变化,使得该方法难以发现异常放电的早期迹象。近年来,基于声纹的变压器工况异常实时监测方法迅速发展,得益于语音识别技术的成熟,利用变压器在运行时产生的物理振动声音来判断其工作状态成为可能^[3,5,8-9]。

目前,大量研究已通过提取采集到的声音信号特征来判断变压器中存在的各类故障。在特征提取方面,主流研究分为时域与频域两种信号特征提取方法:基于时域的提取方式主要包括包络曲线与过零率信息^[10-11],用于判断声音信号瞬间能量或频率的变化;基于频域的提取方式首先使用傅里叶变换^[8],将时域信号转换为频域信号。由于变压器正常工作及常见异常缺陷可能发生的频带非常宽,最高甚至可以达到超声频率。因此频域信号维度非常大,直接分析将占用大量计算资源,必须进行关键频率提取与降维操作,一般通过诸如梅尔倒谱系数(Mel-Frequency Cepstrum Coefficient, MFCC)、线性预测系数(Linear Prediction Coefficient, LPC)以及线性预测倒谱系数(Linear Prediction Cepstrum Coefficient, LPCC)等方式实现关键声纹信息提取^[12-13]。最后,设计声纹分析算法实现异常故障声纹的检测与识别功能。传统的基于信号处理与机器学习的分析方法包括支持向量机(Support Vector Machine, SVM)、隐马尔可夫模型(Hidden Markov Model, HMM)以及高斯混合模型(Gaussian Mixture Model, GMM)等^[14-16]。随着神经网络的出现与发展,后向传播网络(Back Propagation, BP)和深度学习网络(Deep Neural Network, DNN)也被引入到变压器分析领域^[2,7,17],并取得较好的成果。

目前,常见的变压器异常故障声纹如铁心松动、绕组变形等已经被证明可以通过声纹分析获得较好的检测能力^[18]。然而,异常放电故障,尤

其是准确识别轻微的异常放电现象,依然是富有挑战的课题,其主要原因有如下两点。首先,异常放电发生速度极快,通常小于 10 毫秒,而普遍的变压器声纹信号处理需要进行加窗分帧操作,窗体长度通常选取 100 毫秒到 500 毫秒,再对每帧进行快速傅里叶变换获得频域信号。然而,为了捕捉到迅速的异常放电,窗体需要进一步缩小到 20 至 50 毫秒,这极大增加了需要分析的声纹帧数的数量,而诸如隐马尔可夫链、循环神经网络(Recurrent Neural Network, RNN)长短时记忆网络(Long Short-Term Memory, LSTM)都是时序处理网络^[19-20],帧数的增加必然造成计算时间的增加。其次,由于初期的异常放电声音极易被变压器本体振动声音淹没,仅表现为频域上的微弱变化。目前常见的梅尔倒谱系数、线性预测系数缺乏对变压器实际工作的声纹频率的优化,同时当前以循环神经网络为代表的异常故障检测方法难以兼顾瞬时微弱变化。因此,我们不仅需要对异常放电表现特性进行有针对性的特征提取方法设计,还需要研究一种高精度高敏感度的声纹特征变化感知分析方法。

本文首先针对变压器异常放电的特点,即其信号弱且发生迅速,结合关键频率提取与梅尔系数提取思想,设计了一种基于二维化的关键声纹特征提取方法。该方法在保留变压器工频本征特性的基础上,实现了对微弱异常放电的有效提取,并通过高倍率特征维度压缩。其次,本文设计了一种基于卷积神经网络(Convolutional Neural Network, CNN)^[21-22]与视觉转换器(Vision-Transformer, ViT)^[23]的混合神经网络架构。该网络结合了 CNN 的特征提取能力和 ViT 的自注意力机制,实现了对微弱异常放电的高精度发现与识别。最后,本文结合 110 kV 变压器在无载调压试验中的实际放电数据进行了试验及分析,确定了最佳的变压器特征提取方法和识别模型参数。通过这一综合方法,本研究为解决变压器异常放电的弱、快特性提供了有效的解决方案,并在真实数据上取得了令人满意的试验结果。

1 变压器声纹预处理与特征提取

1.1 接触式变压器声纹采集与试验部署

本文采用了接触式拾音器,该拾音器通过磁

吸方式附着于 110 kV 三相三绕组变压器外壳,利用变压器外壳传导变压器内部振动产生的声纹信号。相较于非接触式的枪体麦克风,该方法的优势在于省略了变压器外壳到麦克风之间的空气介质传导过程。这一设计不仅避免了空气介质对微弱声纹信号的削弱现象,同时减少了环境噪音对声音信号的干扰,从而提升了声纹拾音器对微弱信号的感知能力。

为了证明接触式拾音器的有效性,分别使用两种拾音器进行信号采集并进行对比。两种拾音器的声纹信号采样率同为 44.1 kHz,接触式部署于变压器放电试验的近端侧与远端测外壳上,分别位于 1 米和 2 米的高度。非接触式部署于距离变压器 0.5 米与 1 米位置。接触式与非接触式拾音器在跳闸前后的对比声纹时谱图如图 2 所示。对比可知,接触式拾音器在近端 1 米高度的信号 (a) 包络最平稳,2 米高度的信号 (b) 虽然保持了较高的信号强度,但是包络上出现小幅度的波动,这可能与靠近变压器箱体上部,导致采集到了套管等部件的杂音有关。在变压器跳闸前,非接触式拾音器 (c) 相比接触式信号强度明显更小且包络波动更大;当发生跳闸之后,接触式拾音器获取的信号 (a) 强度下降最迅速,而非接触式拾音器

(d) 甚至出现瞬间信号强度大于跳闸前的情况。这是由于跳闸产生的巨大声响在封闭场地环境的空气介质中形成了持久响亮的回声,并被非接触式拾音器持续接收,进一步证明了非接触式变压器更容易受到周围环境的影响。由此,该试验以接触式拾音器近端 1 米高度采集到的数据作为试验数据进行分析。

本试验数据基于 110 kV 三相三绕组变压器无载调压试验。随着电压爬升过程,记录变压器内部产生的声纹信号。图 3 展示了随着电压的爬升,总体声音逐渐增大的现象;图 4 展示了伴随的放电频率逐渐上升的情况。



图 1 变压器拾音器部署图:右上为接触式拾音器,右下为非接触式拾音器

Fig. 1 Transformer voice sensor deployment: contact type-the upper right, non-contact-lower right

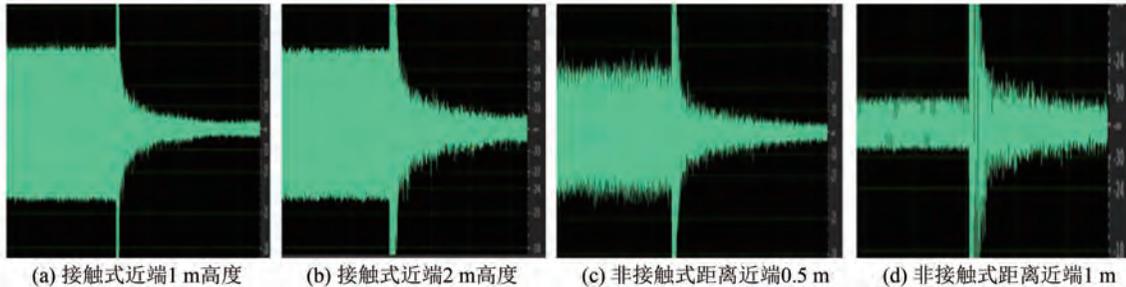


图 2 不同部署位置下拾音器采集的声纹对比

Fig. 2 Comparison of voiceprints collected by pickups at different deployment positions

1.2 变压器声纹样本预处理

为了对采集到的声纹信号进行量化分析,本文采用了一系列预处理步骤,包括切分、分帧、加窗以及离散傅里叶变换。处理的输出将被传递到后续的特征提取模块,流程图如图 5 所示。在本文中,监测的声纹信号每 5 秒进行一次切分,采用交叠分段法进行分帧,帧长设置为 25 毫秒,帧移(步长)为 0.5 帧长,因此每帧约包含 1100 采样点,帧移为 550 采样点。这一处理流程有助于更

好地对声纹信号进行分析和提取有用的特征。

为了减少由于分帧对声纹信号的切分而产生的截断效应,减轻傅里叶变换时可能出现的频率泄露问题,通常会使用汉明窗函数进行帧的加窗操作。汉明窗的表达式如下:

$$w(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{L}\right), & 0 \leq n \leq L - 1 \\ 0, & \text{其他} \end{cases} \quad (1)$$

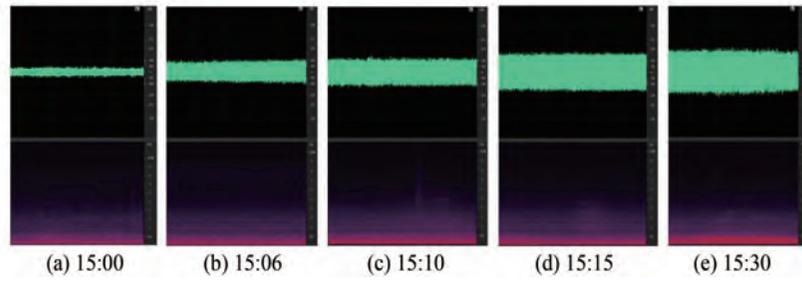


图 3 电压爬升变压器声纹时频谱图变化趋势
 Fig. 3 Spectrum trend when voltage climbs

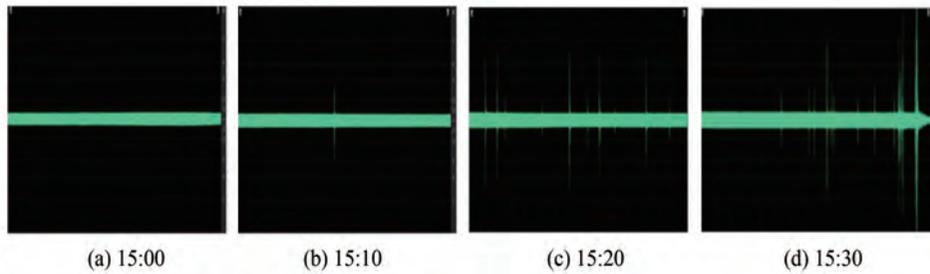


图 4 电压爬升过程中异常放电发生变化趋势
 Fig. 4 Discharge occurrence trend when voltage climbs

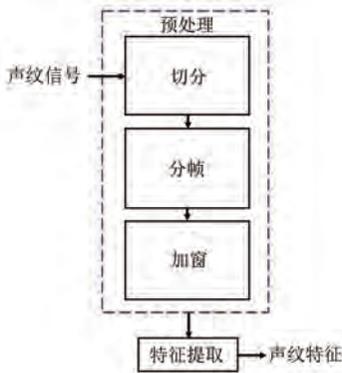


图 5 变压器声纹样本预处理流程框图

Fig. 5 Block diagram of sample preprocessing flow

式中: $w(n)$ 为窗函数在第 n 个采样点的取值; L 为窗长。

在进行每帧数据的快速傅里叶变换 (Fast Fourier Transform, FFT) 操作后, 将时域信号转换为频域信号。如果将获得的每一帧频域信号视为一个多维特征向量, 对全部向量进行拼接操作, 就能够得到一个 $[F, T]$ 的二维特征矩阵, 其中 F 为频域维度, T 为总帧数。在本试验中, 5 秒的声纹数据, 采样率为 44.1 kHz, 步长为 12.5 毫秒。因此, 此时二维特征矩阵的尺度为 $22\ 000 \times 400$ 。由于直接计算将会消耗大量计算资源, 因此需要进一步优化特征提取方法, 以降

低频率特征维度。

1.3 优化的混合频率特征提取

声纹分析目前最常用的特征提取方法是梅尔频率提取法。该方法首先将频域信号转换为梅尔频域, 然后划分成等宽的梅尔频带。这一特征提取方法模拟了人耳对于低频和高频的分辨能力, 体现了人耳对于频率的非线性响应。

频率 f 向梅尔 m 的转换计算式如下:

$$\text{Mel}(f) = 2\ 595 \times \lg\left(1 + \frac{f}{700}\right) \quad (2)$$

式中: $\text{Mel}(f)$ 为频率 f 转换为对应的梅尔频率的变换函数。

从梅尔频率 m 向频域转换计算式如下:

$$\text{Mel}^{-1}(m) = 700 \times (10^{m/2\ 595} - 1) \quad (3)$$

式中: $\text{Mel}^{-1}(m)$ 为梅尔频率 m 重新转换成对应的频率的逆变换函数。

然而, 变压器的工作频率主要集中在 1 000 Hz 以下的 50 Hz 倍频。变压器大部分异常缺陷发生都会引起这些频率的明显变化, 因此在频率提取中保留该频带的分辨率尤为重要。如表 1 所示, 梅尔滤波器组的峰值频率很难与 50 Hz 倍频对齐, 因此使用梅尔滤波器组可能导致 1 000 Hz 以下关键频率特征的损失。

表 1 不同数量的三角滤波器峰值频率 (Hz)

Tab. 1 Peak frequencies of triangular filters with different amounts (Hz)

滤波器 编号	滤波器数量				
	20	22	24	26	28
1	66	60	55	51	47
2	139	126	115	106	98
3	218	197	179	228	152
4	306	275	249	295	210
5	401	359	324	368	272
6	506	450	405	446	338
7	620	550	493	530	408

因此,本文提出将梅尔频率与关键频率提取相结合的方法:

(1) 计算使用 400 个三角滤波器的梅尔滤波器组覆盖声纹信号全频率时每个滤波器峰值频率;(2) 仅计算峰值在 1 000 Hz 以上的梅尔滤波器组特征值;(3) 对 1 000 Hz 及以下频段使用 20 Hz 频带交叠最大值采样法,即从 1 Hz 起,每 20 Hz 频带取其最大值作为采样特征,频带间交叠 1/2,共获得 100 维特征向量;(4) 该 100 维特征与获得的梅尔频率特征拼接,从而得到一个混合频率特征。

由(1)获得的滤波器组中第 85 个滤波器峰值为 985.3 Hz,第 86 个滤波器峰值为 1 002.8 Hz,则(2)中仅计算第 86 至第 400 共 315 个三角滤波器的梅尔频率特征值。(3)中获得 1 000 Hz 以下 100 维特征值,(4)中可知,5 秒声纹信号经过该方法可获得尺度为 415×400 的二维特征矩阵。该混合特征提取方法在保证变压器工频分辨率同时,极大压缩了频率维度,提升了后续计算资源利用效率。

2 基于混合神经网络的变压器异常放电识别方法

异常放电与变压器工作振动相比能量较轻微,经常不能在时域谱线中表现出来。传统的包络突变检测等方法难以保证检测的召回率。同时,异常放电发生的时间较短,存在突发性,与时序上下文信号关联性较少。因此,常见的基于时序的识别方法,如 HMM、RNN 和 LSTM 等,往往不能很好地发现异常放电现象的发生。此外,时序分析方法处理时间与声纹样本切片长度呈正相

关,并呈现指数上升趋势,不适合分析较长的声纹样本切片。相反,对声纹样本进行二维化分析既兼顾了瞬时频率表征与时序之间的关系,又得益于卷积神经网络对于特征矩阵的出色处理能力,实现了高效的特征语义提取。Transformer 架构的识别网络相较于传统的直接使用全连接层分类网络获得可学习分类权重的方法,使用了一种级联多个编码器模块结构完成对语义特征进一步分析。其中,每个编码器模块中包含一个多头注意力模块,利用自注意力机制学习语义特征中各个序列化的局部特征与全局语义之间的关联程度。这进一步实现了在较为稀疏的二维特征矩阵中定位到微弱的异常放电特征,从而实现了高精度的识别表现。完整网络架构图如图 6 所示。

2.1 基于卷积神经网络的二维特征提取

首先,需要对声纹特征提取阶段获得的二维特征矩阵进行尺度变换,将矩阵的长宽从 415×400 扩充为 448×448,然后转换为可应用于卷积神经网络的三维特征张量 448×448×1。在设计特征语义提取网络时,需要解决梯度消失和梯度爆炸的问题,同时防止微弱的异常放电视觉特征随着卷积层数的增加而消失。考虑到这些问题,可采用残差网络架构。鉴于整体运算速度的考虑,可以选择使用 ResNet18 网络作为特征语义提取网络。该网络架构如图 7 所示。

首先,ResNet18 的特征输入通道数从 3 改为 1,以适配二维声纹特征的通道数。之后,遵循原版 ViT 代码中的 ResNet50 搭建细节 (https://github.com/google-research/vision_transformer),对 ResNet18 进行相应修改,包括:(1) 将原版 ResNet18 中全部的 BatchNorm 操作替换为 GroupNorm 操作;(2) 网络中所有卷积层使用 PyTorch 中的标准化权重的二维卷积 (nn. StdConv2d) 替换为普通二维卷积 (nn. Conv2d)。经过 ResNet18 网络后,将得到一个 14×14×512 的特征图,作为后续 Vision-Transformer 的输入。

2.2 基于 Vision-Transformer 的异常放电识别

由于变压器的异常放电通常会影响到频谱信息的多个频段,即放电语义与特征图中的多个局部区域存在较强的关联性,且不同程度的放电影响力也不尽相同。因此,相较于直接使用全连接层设计分类器,Vision-Transformer 的自注意力机制

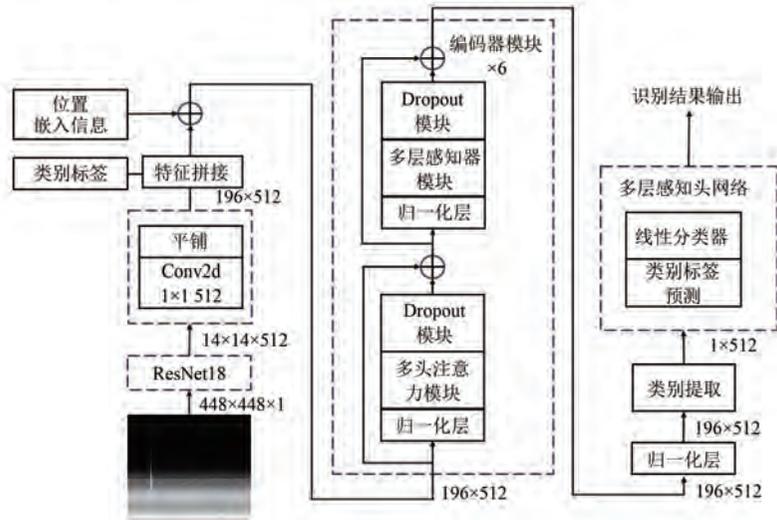


图 6 混合网络架构图

Fig. 6 Hybrid network structure

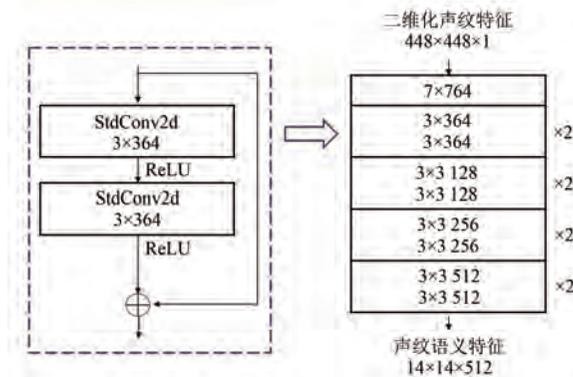


图 7 修改后的 ResNet18 网络

Fig. 7 Modified ResNet18 network

能够有效实现语义与特征图序列的关联对应,从而有效提升识别效果。

CNN+ViT 混合网络架构如图 8 所示。考虑到变压器声纹特征相对简单,且需要尽可能减少网络参数量以提高处理速度,将 ViT 网络中的编码器模块的层数从原文 Base 结构的 12 层缩减到 6 层。在 3.3 节中,我们进行了缩减层数前后的网络识别效果对比,试验证明该策略能够有效减少神经网络收敛所需的时间,并且保留 ViT 网络在自注意力上的优势。

3 变压器异常放电识别试验

3.1 数据与试验平台参数

本文的试验数据来源于 110 kV 三相三绕组变压器无载调压放电试验中多路拾音器采集的数

据。人工从数据中筛选出了 240 条放电样本和 1 200 条未放电样本,其中 80% 被随机选取作为训练样本,剩余 20% 作为测试样本。每个样本的长度为 5 秒,采样频率为 44.1 kHz,经过快速傅里叶变换后,频域信号仅保留前 22 kHz 用于后续分析。图 8 展示了三种不同放电强度案例的部分典型样本。在时域中,(b)与(c)并没有明显的包络变化,而(c)在时频谱线中也仅出现了几乎难以察觉的变化,即放电信号所携带的能量极低。因此,当放电瞬间的信号能量与非放电瞬间信号能量比例越相近时,该异常放电现象越微弱。

本次试验平台使用 Linux 18.04 操作系统,CPU 型号 Intel i9 9900X,内存 32 GB,GPU NVIDIA RTX Titan 24 GB。

3.2 声纹数据特征提取方法对比试验

为了选取最佳的声纹特征提取方法,本部分使用了多种不同的声纹特征提取方法进行比较,包括梅尔频率系数以及本文提出的二维声纹特征提取方法。为了找到最佳的二维声纹特征提取方法,分别使用了不同数量的梅尔滤波器组进行试验。在该试验中,所有单帧特征都会通过时序合并的方式转换为二维声纹特征。为了保证使用统一结构的神经网络,所有二维声纹特征都将被缩放为尺度为 448x448x1 的特征张量,作为识别神经网络的输入。

本次试验的识别对比结果如表 2 所示。通过识别率的比较可以发现,所提出的混合特征提取

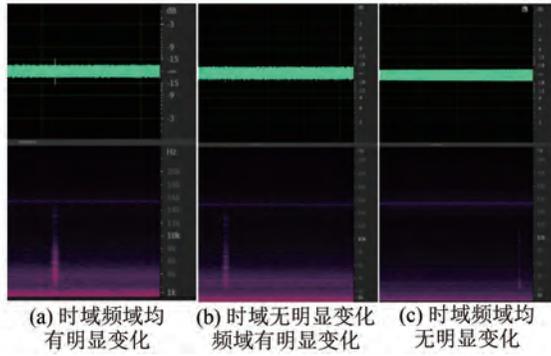


图 8 异常放电样本时域频域谱图: (a) 时域频域均有明显变化, (b) 时域无明显变化频域有明显变化, (c) 时域频域均无明显变化

Fig. 8 The time-domain and frequency-domain spectrum of abnormal discharge samples; (a) significant changes in the time-domain and frequency-domain, (b) no obvious changes in the time-domain and obvious changes in the frequency-domain, (c) no obvious changes in the time-domain and frequency-domain

方法整体上要优于直接使用梅尔频率系数进行特征提取方法。同时,无论是梅尔频率系数还是所提特征提取方法,增加滤波器组数量时都会提升效果,但滤波器组的增加明显会增加运算时间,因此需要在滤波器组数量与识别表现之间取得平衡。相较于传统梅尔频率特征提取方法,识别率提升约 8.3%。相较于全频段 50 Hz 最大值特征提取,识别用时减少约 0.18 秒/样本。

表 2 不同特征提取方法识别用时及识别率对比

Tab. 2 Comparison of recognition time and accuracy of different feature extraction methods

特征提取方法	识别用时(全测试集)/秒	识别率/%
梅尔滤波器,数量 200	271	81.60
梅尔滤波器,数量 400	342	85.07
全频段 50 Hz 最大值提取法	417	89.58
全频段 50 Hz 交叠最大值提取法	429	89.93
混合特征提取方法,数量 200	295	86.46
混合特征提取方法,数量 400	377	89.93

3.3 异常放电声纹特征识别方法对比试验

为了对比不同识别方法对异常放电声纹信号的识别表现,本部分使用多种常见识别方法进行识别结果比较,其中包括 LSTM、ResNet18、ResNet50 以及本文提出的 CNN+ViT 方法。特征

提取方法使用本文所提混合特征提取方法,识别结果对比如表 3 所示。

表 3 不同识别方法识别用时与识别率对比

Tab. 3 Comparison of recognition time and accuracy of different recognition methods

识别方法	识别用时(全测试集)/秒	识别率/%
LSTM	672	71.53
ResNet18	197	81.25
ResNet50	331	85.42
ViT	354	85.76
CNN+ViT	377	89.93

由表 3 可知,所提方法相较于 LSTM 网络,不仅在识别用时上减少 295 秒(约 1.02 秒/样本),而且在识别率上提升 18.4%,这说明异常放电与时序上下文关联度不高,以 LSTM 为代表的时序识别网络不适合于该任务。使用 ResNet18 与 ResNet50 网络后,识别率上明显提升。而单纯使用 ViT 网络,由于缺乏深度卷积神经网络的语义提取能力,其提升效果极其有限。所提的 CNN+ViT 方法识别率表现得到进一步提升,相较于 ResNet50 识别用时仅增加 46 秒(约 0.16 秒/样本),识别效果提升约 4.5%,相较于 ViT 网络识别效果提升约 4.2%,证明基于 CNN 网络的语义特征提取能够较好的获取二维化特征中的异常放电特征。

为进一步说明 2.2 节中本方法模型构建时减少编码器模块层数带来的效果提升,本部分进行了修改前后的模型表现对比,对比结果如表 4 所示。

表 4 编码器模块层数缩减前后模型表现对比

Tab. 4 Comparison of the performance of the model before and after the reduction of the number of encoder module

识别率/%	迭代次数 (Epoch)							
	2	4	6	8	10	12	16	24
CNN+ViT(12层)	51.7	68.8	78.5	83.3	86.7	87.1	89.6	89.2
CNN+ViT(6层)	53.5	74.0	81.3	86.5	89.9	89.6	89.2	88.9

由上表可知,使用 6 层编码器模块的模型在第 10 次迭代时达到了最佳效果,而使用 12 层编码器模块的模型需要到第 16 次迭代才能达到相似的识别率。由此可见,在本任务中,适当减少编码器层数不仅能够减少模型的计算参数量,还有助于提高模型达到最佳效果的收敛速度。需要指

出的是,两个模型在达到最佳表现后均出现了下降的趋势,这可以理解为由于数据总量及数据特征丰富度的局限性导致的过拟合现象。

4 结语

本文通过对采集的 110 kV 三相三绕组变压器无载调压过程中异常放电声纹数据进行分析,提出了一种优化的二维声纹特征提取方式,并构建了一个基于卷积神经网络与 Transformer 架构网络的混合异常放电特征识别方法。通过与现有特征提取方法及识别方法在识别用时与识别率上的对比,所提出的混合特征提取方法相较于传统梅尔频率特征提取方法,识别率提升约 8.3%。与时序神经网络 LSTM 相比,在识别用时及识别率上分别提升 1.02 秒/样本和 18.2%。相对于卷积神经网络 ResNet50,识别率提升 4.5%,比 ViT 网络识别率提升 4.2%。试验证明所提出的特征提取方法与识别方法在保证识别速度的同时,提高了异常放电声纹的识别准确性。

参考文献

- [1] 刘云鹏,王博闻,岳浩天,等. 基于 50 Hz 倍频倒谱系数与门控循环单元的变压器偏磁声纹识别[J]. 中国电机工程学报, 2020, 40(14): 4681-4694+4746.
LIU Y P, WANG B W, YUE H T, et al. Identification of transformer bias voiceprint based on 50 Hz frequency multiplication cepstrum coefficients and gated recurrent unit [J]. Proceedings of the CSEE, 2020, 40(14): 4681-4694+4746.
- [2] 于达,张玮,王辉. 基于 LSTM 神经网络的油浸式变压器异常声纹诊断方法研究[J]. 智慧电力, 2023, 51(2):45-52.
YU D, ZHANG W, WANG H. Abnormal voiceprint diagnosis method of oil-immersed transformer based on LSTM neural network [J]. Smart Power, 2023, 51(2): 45-52.
- [3] 汲胜昌,师榆航,张凡,等. 电力变压器振动与噪声及其控制措施研究现状与展望[J]. 高压电器, 2019, 55(11): 1-17.
JI S C, SHI Y H, ZHANG F, et al. Review on vibration and noise of power transformer and its control measures [J]. High Voltage Apparatus, 2019, 55(11): 1-17.
- [4] 李刚,于长海,刘云鹏,等. 电力变压器故障预测与健康管理的挑战与展望[J]. 电力系统自动化, 2017, 41(23): 156-167.
LI G, YU C H, LIU Y P, et al. Challenges and prospects of fault prognostic and health management for power transformer [J]. Automation of Electric Power Systems, 2017, 41(23): 156-167.
- [5] 刘云鹏,罗世豪,王博闻,等. 基于 Mel 时频谱-卷积神经网络的变压器铁芯夹件松动故障声纹模式识别[J]. 华北电力大学学报(自然科学版), 2020, 47(6): 52-60+67.
LIU Y P, LUO S H, WANG B W, et al. Voiceprint recognition of transformer core clamp looseness fault by Mel-spectrum and convolutional neural network [J]. Journal of North China Electric Power University (Natural Science Edition), 2020, 47(6): 52-60+67.
- [6] 范文杰,张志斌,夏昌杰,等. 多手段测量下的油纸绝缘针板电极局部放电的演化规律分析[J]. 高电压技术, 2022, 48(3): 914-927.
FAN W J, ZHANG Z B, XIA C J, et al. Evolution law of partial discharge of oil-paper insulated needle-plate electrode based on multi-method measurement [J]. High Voltage Engineering, 2022, 48(3): 914-927.
- [7] 施胜丹,黄金军,朱霄珣,等. 基于声纹 SDP-CNN 的变压器局部放电模式识别[J]. 电力信息与通信技术, 2022, 20(10): 105-112.
SHI S D, HUANG J J, ZHU X X, et al. Partial discharge pattern recognition on transformer based on voiceprint SDP-CNN [J]. Electric Power Information and Communication Technology, 2022, 20(10): 105-112.
- [8] 张可,杨可军,黄文礼,等. 基于声纹识别的变压器工况检测方法验证系统[J]. 计算技术与自动化, 2022, 41(1): 1-6.
ZHANG K, YANG K J, HUANG W L, et al. Detection algorithm and verification system of transformer working condition based on voiceprint recognition [J]. Computing Technology and Automation, 2022, 41(1): 1-6.
- [9] 王啸峰. 浅谈智能变电站声纹监测技术的应用[J]. 机电信息, 2022, 6: 12-16.
WANG X F. Talking about the application of voiceprint monitoring technology in intelligent substation [J]. Mechanical and Electrical

- Information, 2022, 6: 12-16.
- [10] 张重远, 罗世豪, 岳浩天, 等. 基于 Mel 时频谱-卷积神经网络的变压器铁芯声纹模式识别方法 [J]. 高电压技术, 2020, 46(2): 413-422.
 ZHANG C Y, LUO S H, YUE H T, et al. Pattern recognition of acoustic signals of transformer core based on Mel-spectrum and CNN [J]. High Voltage Apparatus, 2020, 46(2): 413-422.
- [11] 杜一明. 基于声信号的变压器故障诊断系统研究 [D]. 武汉: 华中科技大学, 2013.
 DU Y M. Research on fault diagnosis system for power transformer based on audio signal [D]. Wuhan: Huazhong University of Science and Technology, 2013.
- [12] 王丰华, 王邵菁, 陈颂, 等. 基于改进 MFCC 和 VQ 的变压器声纹识别模型 [J]. 中国电机工程学报, 2017, 37(5): 1535-1543.
 WANG F H, WANG S J, CHEN S, et al. Voiceprint recognition model of power transformers based on improved MFCC and VQ [J]. Proceedings of the CSEE, 2017, 37(5): 1535-1543.
- [13] 余建潮, 张瑞林. 基于 MFCC 和 LPCC 的说话人识别 [J]. 计算机工程与设计, 2009, 30(5): 1189-1191.
 YU J C, ZHANG R L. Speaker recognition method using MFCC and LPCC features [J]. Computer Engineering and Design, 2009, 30(5): 1189-1191.
- [14] 赵康. 改进 SVM 结合决策树的情感语音识别 [J]. 信息技术, 2020, 44(8): 17-22.
 ZHAO K. Speech emotion recognition using decision tree and improved SVM [J]. Information Technology, 2020, 44(8): 17-22.
- [15] RABINER L R. A tutorial on hidden Markov models and selected applications in speech recognition [J]. Proceeding of the IEEE, 1989, 77(2): 258-286.
- [16] 周梦茜, 唐志国, 王泽瑞, 等. 基于声纹识别系统的局部放电超声信号识别研究 [J]. 高压电器, 2022, 58(9): 127-133.
 ZHOU M X, TANG Z G, WANG Z R, et al. Study on ultrasonic signal recognition of partial discharge based on voiceprint recognition system [J]. High Voltage Apparatus, 2022, 58(9): 127-133.
- [17] 余长厅, 黎大健, 汲胜昌, 等. 基于振动噪声及 BP 神经网络的变压器故障诊断方法研究 [J]. 高压电器, 2020, 56(6): 256-261.
 YU C T, LI D J, JI S C, et al. Research on transformer fault diagnosis method based on vibration noise and BP neural network [J]. High Voltage Apparatus, 2020, 56(6): 256-261.
- [18] 张晨晨, 丁国成, 李坚, 等. 基于人工智能和数据驱动的电力变压器声纹识别探索与应用 [J]. 电器与能效管理技术, 2021, 11: 85-89.
 ZHANG C C, DING G C, LI J, et al. Exploration and application of voiceprint recognition of power transformer based on artificial intelligence and data driving [J]. Electrical & Energy Management Technology, 2021, 11: 85-89.
- [19] BERINGER N, GRAVES A, SCHIEL F, et al. Classifying unprompted speech by retraining LSTM nets [C] // Artificial Neural Networks: Biological Inspirations-icann, International Conference, Warsaw, 2005.
- [20] BAKKER B. Reinforcement learning with long short-term memory [J]. Advances in Neural Information Processing Systems, 2001, 14: 1475-1482.
- [21] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
- [22] 何萍, 李勇, 陈寿龙, 等. 基于变压器声纹 Mel 语谱图-ResNet 的铁心松动故障诊断 [J]. 电机与控制应用, 2022, 49(9): 75-80.
 HE P, LI Y, CHEN S L, et al. Fault diagnosis of iron core looseness based on Mel spectrogram-ResNet with transformer voiceprint [J]. Electric Machines & Control Application, 2022, 49(9): 75-80.
- [23] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale [C] // International Conference on Learning Representations, 2021.

收稿日期:2023-07-27

收到修改稿日期:2023-10-23

作者简介:

童 吻(1991-),男,硕士,主要从事于变电站智能运检技术研究,380493337@qq.com;

* 通信作者:黄文礼(1987-),男,硕士,高级工程师,主要从事音视频人工智能技术在电力系统中的应用研究,491930484@qq.com。

A Recognition Method Based on Two-Dimensional Voiceprint Feature Extraction Method and Hybrid Neural Network for Contact-Collected Abnormal Discharge Voiceprint of 110 kV Three-Phase Three-Winding Power Transformer in No-Load Voltage Regulation

TONG Yang, HUANG Wenli^{*}, LI Lei, YAN Yuqing

(Anhui NARI Jiyuan Power Grid Technology Co., Ltd., Hefei 230088, China)

Key words: abnormal discharge of transformer; voiceprint recognition; voiceprint feature extraction; hybrid neural network

Power transformer is a vital power facility in the process of transmission, transformation and distribution, and ensuring the normal operation of power transformer is an important part of ensuring the security of power transmission and supply. With the passage of transformer working time, it is inevitable that the transformer working conditions will be affected due to operating load operation, aging equipment parts, turbid insulation oil, etc., such as loose core, winding deformation, short circuit between turns, DC bias, abnormal discharge and other defects. Among them, abnormal discharge is one of the common transformer defects cases with a high-risk factor.

In this paper, aiming at the characteristics of weak and fast abnormal discharge of transformers, combined with the idea of key frequency extraction and Mel cepstral coefficient, a key voiceprint feature extraction method based on two-dimensional design is designed, which realizes the effective extraction of weak abnormal discharge and high-rate feature dimension compression on the basis of retaining the intrinsic characteristics of transformer power frequency. Secondly, this paper designs a transformer abnormal discharge recognition model based on a hybrid neural network architecture of convolutional neural network and Vision-Transformer (ViT), and uses the self-attention mechanism of the network to achieve high-precision detection and

recognition of weak abnormal discharge. Finally, this paper combines the real discharge data of real transformers to conduct experiments and analysis to determine the best transformer feature extraction method and identify model parameters.

By comparing with the effect of Mel coefficient, the experimental results show that the proposed feature extraction method can take into account the transformer power frequency and high frequency feature extraction, obtain better recognition accuracy and reduce the calculation time of feature extraction.

The proposed hybrid neural network uses the redesigned ResNet18 network as well as the ViT network to detect weak discharge features from the voiceprint signal and achieve accurate recognition. The results are shown in Tab.1.

Tab. 1 Comparison of recognition time and accuracy of different recognition methods

Methods	Time (all test set)/s	Accuracy/%
LSTM	672	71.53
ResNet18	197	81.25
ResNet50	331	85.42
CNN+ViT	377	89.93

By comparing with the existing feature extraction methods and recognition methods in recognition time and recognition rate, the proposed feature extraction method and recognition method take into account the recognition speed and recognition accuracy.