

# 结合连续小波与 SCF-AE 网络的牵引变压器 异常声纹检测

李喆

(国能朔黄铁路发展有限责任公司, 河北 沧州 062350)

**摘要:** 牵引变压器是保障牵引供电系统正常运行的关键设备; 既有变压器声纹检测方法基于故障样本驱动, 在声纹样本稀缺的重载铁路牵引变压器状态检测场景下适用性受限; 针对该问题, 提出结合连续小波与 SCF-AE 网络的牵引变压器异常声纹检测方法; 采用连续小波变换, 获取了牵引变压器声纹的时频特征图; 通过在自编码器中引入空间通道融合卷积, 提升网络的特征学习能力的同时降低了其计算复杂度; SCF-AE 网络在训练阶段对牵引变压器正常声纹样本进行了压缩重构; 在检测阶段, 基于该网络, 通过分析输入样本的重构误差是否超过阈值来判断是否存在异常; 经试验测试, SCF-AE 在数据集上实现了 88.17% 的精确率和 100% 的召回率, 对单个样本的平均推理时长仅为 13.3 ms, 对比当前既有方法其准确性和计算复杂度均具有优势。

**关键词:** 牵引变压器; 声纹; 异常检测; 连续小波变换; 自编码神经网络; 无监督学习

## Anomaly Acoustic Fingerprint Detection for Traction Transformers Integrated Continuous Wavelet Transform with SCF-AE Network

LI Zhe

(China Energy Shuohuang Railway Development Co., Ltd., Cangzhou 062350, China)

**Abstract:** Traction transformers are regarded as critical devices for ensuring the reliable operation of traction power supply systems. Based on fault sample-driven approaches, Existing acoustic fingerprint detection methods for traction transformers have a limited adaptability for monitoring the status of heavy-haul railway traction transformers with scarce acoustic fingerprint samples. To address this issue, an abnormal acoustic fingerprint detection method for traction transformers is proposed, which integrates continuous wavelet transform with a spatial-channel fusion autoencoder (SCF-AE) network. A continuous wavelet transform is used to achieve the time-frequency feature map of transformer acoustic fingerprints. A spatial-channel fusion convolution is introduced into the autoencoder, which enhances the feature learning capability of the network while reducing its computational complexity. During the training phase, normal acoustic fingerprints for traction transformers are compressed and reconstructed in the SCF-AE network. During the detection phase, anomalies are identified by determining whether the reconstruction error of an input sample exceeds the threshold. Through test on the dataset, the SCF-AE method achieves a precision of 88.17% and a recall of 100%, with an average inference time of only 13.3 ms for a single sample, which is superior in accuracy and computational efficiency to existing methods.

**Keywords:** traction transformer; sound pattern; anomaly detection; continuous wavelet transform; autoencoder neural network; unsupervised learning

## 0 引言

朔黄铁路是我国西煤东运第二通道“神黄铁路通道”的关键组成部分<sup>[1]</sup>, 围绕其牵引变压器开展状态检

测, 从而维护其牵引供电系统的正常安全运行, 对于保障我国能源安全和经济发展有重要意义。朔黄铁路作为重载货运铁路, 其牵引负荷的冲击相比常规电气化铁路更大<sup>[2]</sup>, 如图 1 所示为朔黄铁路某牵引所实测馈线电流

收稿日期:2025-05-22; 修回日期:2025-07-11。

基金项目:国能朔黄铁路发展有限责任公司科技创新项目(SHTL-24-42)。

作者简介:李喆(1974-),男,硕士,教授级高级工程师。

引用格式:李喆. 结合连续小波与 SCF-AE 网络的牵引变压器异常声纹检测[J]. 计算机测量与控制, 2026, 34(4): 71-78.

幅值变化曲线,可以看到在没有列车通行时电流幅值极低,而当货运列车通过时电流幅值激增,具有极强冲击性。电流增大会使得绕组所受电动力增大,长期作用下易导致变压器绕组出现累积性的应力疲劳<sup>[3]</sup>,且由于铁路机车采用 AC/DC 整流方式,导致负荷电流的谐波成分较大,产生额外的变压器绕组和铁芯损耗,加剧变压器老化<sup>[4]</sup>、松动等机械异常<sup>[5]</sup>导致其抗短路性能严重劣化,进而引发潜在的变压器故障失效并危害整个牵引供电系统的正常运行。

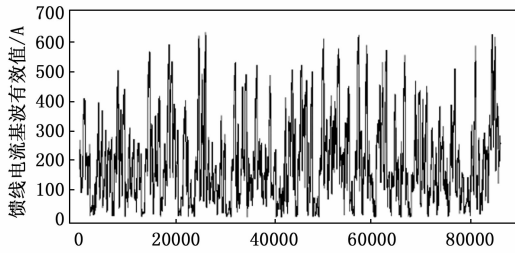


图 1 朔黄铁路某牵引所实测馈线电流数据

传统的牵引变压器状态在线检测方法主要是油色谱分析 (DGA, dissolved gas analysis), 该类方法根据油中溶解的各类气体的比例,并结合不同的智能诊断算法<sup>[6-]</sup>实现对变压器内部状态的精确识别。然而, DGA 主要针对变压器的热故障和电故障,无法非常有效地对变压器机械异常进行检测<sup>[9]</sup>。为实现更全面的牵引变压器状态检测,从而更好地维护朔黄铁路牵引供电系统的安全可靠运行,有必要开展针对牵引变压器机械异常的状态检测方法研究。

受到磁致伸缩等效应的影响,变压器在运行过程中会发生振动并产生相应的声纹信号<sup>[10]</sup>,这些信号中包含大量的设备状态信息<sup>[11]</sup>。通过对声纹信号的分析,可以对包括机械异常在内的多种牵引变压器异常状态进行检测<sup>[12]</sup>。同时,在声纹信号采集过程中传感器无需直接接触变压器,因此,基于声纹识别的变压器状态检测方法具有安全性高、设备装设难度低的巨大优势,在近期受到了大量研究者的关注。上海交通大学王丰华等人首先提取变压器声纹的梅尔倒谱系数 (MFCC, mel frequency cepstrum coefficient),并利用主成分分析方法降低 MFCC 的特征维度,最后通过矢量量化算法实现对变压器铁芯松动故障的准确识别<sup>[13]</sup>。华北电力大学张重远等人提取变压器声纹的 Mel 时频谱图特征,并基于卷积神经网络 (CNN, convolutional neural network) 实现对变压器直流偏磁的准确识别<sup>[14]</sup>。河海大学马宏忠等提出基于多特征声纹图谱的变压器在线检测方法,通过采用多种特征提取手段从原始声纹数据中抽取关键特征,并通过训练 CNN 实现了对变压器绕组松

动的准确检测<sup>[15]</sup>。中国电力科学研究院齐子豪等人通过深度学习及振动熵等方法提取变压器声纹的多粒度特征,并基于迁移学习方法实现准确的变压器故障诊断,通过多角度特征提取实现对噪声的鲁棒<sup>[16]</sup>。

当前研究已基于声纹识别技术实现了准确的电力变压器声纹状态检测。然而,在重载铁路牵引变压器状态检测场景中,变压器故障声纹样本相对稀缺,既有的有监督学习方法在故障样本稀缺时易出现过拟合问题,导致其状态检测性能下降<sup>[17]</sup>。除此之外,牵引变压器所处环境噪声干扰更强,以短时傅里叶变换为基础的传统声纹特征提取方法时频分辨率低,难以挖掘牵引变压器声纹的关键特征<sup>[18-19]</sup>,导致变压器状态检测模型的准确性进一步受限。

针对上述问题,提出结合连续小波与空间通道融合自编码网络 (SCF-AE, spatial channel fusion-auto encoder) 的牵引变压器声纹异常检测方法。首先,通过连续小波变换对原始牵引变压器声纹信号进行特征提取,获得抗噪声干扰性能更强、分辨率更高的声纹时频特征;其次,在传统自编码网络中引入空间通道融合 (SCF, spatial channel fusion) 卷积<sup>[20]</sup>,降低网络参数数量的同时提升其特征挖掘能力,提出的 SCF-AE 网络无需异常样本训练即可实现牵引变压器异常声纹检测;最后,在朔黄铁路牵引所采集的牵引变压器声纹数据上对提出的声纹异常检测方法性能进行验证,并与其他不同的故障诊断方法进行了对比,实验结果以及后续的分析证明了所提方法的有效性。

## 1 牵引变压器声纹状态检测方法

### 1.1 基于连续小波变换的声纹特征提取

受磁致伸缩以及电动力作用,变压器运行过程中绕组和铁芯会振动并产生声信号,这些信号中包含大量的设备状态信息。通过分析声纹信号可实现准确的变压器机械异常检测。然而,牵引变压器所处环境噪声干扰强,传统声纹特征提取方法时频分辨率低,难以挖掘牵引变压器声纹的关键特征,导致变压器状态检测模型的准确性进一步受限。

连续小波变换 (CWT, continuous wavelet transform) 是一种对信号的时频分析方法,其首先对母小波进行不同尺度的膨胀、伸缩以及平移得到子小波,并随后通过子小波在时频域上分解信号,具有良好的时频分辨特性以及抗噪声性能。

为解决传统声纹特征提取方法的不足,采用基于 CWT 的牵引变压器声纹特征提取方法,针对变压器声纹信号, CWT 的计算公式如式 (1) 所示:

$$CWT(a, b) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{a}} \psi^* \left( \frac{t-b}{a} \right) dt \quad (1)$$

式中， $t$  变量为时间， $x(t)$  是采集到的变压器声信号， $\Psi(t)$  是母小波函数， $\Psi^*(t)$  是母小波函数的复共轭，而  $a$  和  $b$  分别是 CWT 的尺度因子和平移因子。

以短时傅里叶变换 (STFT, short time fourier transform) 和梅尔时频谱图 (Mel, mel spectrogram) 为代表的传统方法时频分辨率受限于滑动窗口大小，时频分辨率不足。对比之下，CWT 具有更高的时频分辨率，可精确描述牵引变压器声纹信号其频率成分随时间的变化趋势，从而为后续的状态检测模型提供更全面的特征信息。在分析过程中，考虑采用 Morlet 小波母函数，其表达式为，其中  $f_c$  为中心频率，默认为 1 Hz：

$$\psi(t) = \pi^{-1/4} e^{j2\pi f_c t} e^{-t^2/2} \quad (2)$$

牵引变压器音频信号中包含大量的 50 Hz 及 100 Hz 倍频的正弦波成分。Morlet 小波对于信号中正弦或近似正弦成分敏感，具有良好的局部时频特征提取能力，在基于振动信号的变压器绕组故障诊断中被成功应用<sup>[21]</sup>，因此采用 Morlet 小波母函数可有效地提取牵引变压器音频中的特征。

在完成小波母函数选取后，需确定尺度因子和位移因子参数；由于变压器音频有效的频率分量位于 10~2 000 Hz 内<sup>[8]</sup>，而在朔黄铁路牵引所内实测到的变压器音频数据采样率为 16 kHz，而小波尺度  $a$  的计算如式 (3) 所示：

$$a = \frac{f_c}{f\Delta t} \quad (3)$$

其中： $\Delta t$  为信号的采样间隔， $f$  则为需要考虑的频率范围，分别带入 10 Hz 和 2 000 Hz 分量，可根据式 (3) 计算得出  $a \in [8, 1\ 600]$ ，由于变压器音频为窄带信号，故采用线性尺度。小波变换基于 pywt 库实现，平移因子默认为 1。

对现场采集的牵引变压器声纹信号的 STFT 时频谱图及 CWT 时频谱图如图 2 所示，从图可见，短时傅里叶变换无法精确描述牵引变压器声纹在某一时刻的频率成分，导致整个特征图时频分辨率差，难以辨识出有效特征；连续小波时频谱图可以更好的建立信号中不同频率成分随时间的变化，可以从图 2 中观察到牵引变压器声纹中的不同频率成分的变化存在一定规律。良好的时频分辨率有助于状态检测模型从中挖掘关键特征，为后续准确的牵引变压器状态检测提供支撑。

## 1.2 SCF-AE 网络

### 1.2.1 卷积自编码网络

对比电力变压器，可用于训练模型的故障样本收集更加困难且数量稀少，难以支撑神经网络的训练<sup>[22]</sup>，使用过少的故障样本训练极易导致神经网络出现过拟合问题，使得声纹状态检测的精度急剧下降。为解决故障

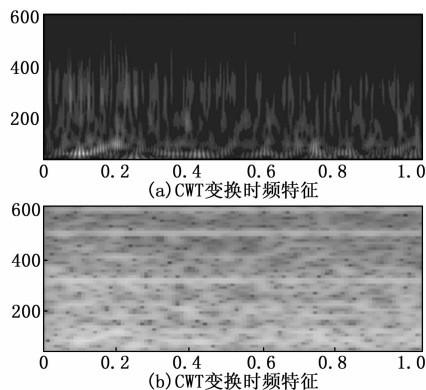


图 2 变压器声纹的 STFT 和 CWT 变换结果

样本数量稀少导致的神经网络性能受限问题，提出一种基于改进卷积自编码神经网络的牵引变压器声纹状态检测模型。所提出的方法仅通过正常声纹样本训练，实现准确的牵引变压器异常状态检测。

卷积自编码神经网络<sup>[23]</sup>的结构如图 3 所示，包括编码器及解码器两部分。其中，编码器的功能是压缩输入的声纹时频谱图，得到压缩编码；随后压缩编码将被输入到解码器中，解码器将对压缩编码进行扩张重构，得到重构输出。

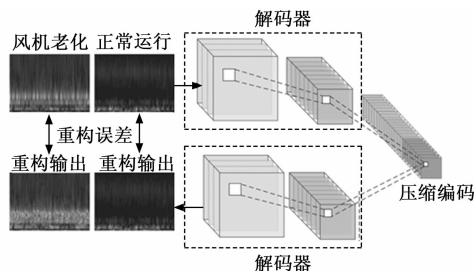


图 3 卷积自编码网络结构

其中，编码器的功能是压缩输入的声纹时频谱图，得到压缩编码；随后压缩编码将被输入到解码器中，解码器将对压缩编码进行扩张重构，得到重构输出。

在训练阶段，首先对牵引变压器原始声信号进行 CWT 变换，得到 CWT 时频谱图；随后在训练过程中，仅在正常牵引变压器声纹数据集上训练自编码网络，训练目标是 minimized 输入声纹时频谱图和重构声纹时频谱图之间的误差，训练目标损失  $L$  如式 (4) 所示：

$$L = L_{MSE} + L_{SSIM} \quad (4)$$

式中， $L_{MSE}$  为自编码网络的重构输出和输入之间的均方误差，而  $L_{SSIM}$  为自编码网络的重构输出和输入之间的结构相似性误差。

$L_{MSE}$  计算方式如式 (5) 所示：

$$L_{MSE} = \frac{1}{h \cdot w} \sum_{i=1}^{h \cdot w} (y_i - x_i)^2 \quad (5)$$

式中， $h$  和  $w$  分别代表重构输出的高度和宽度，而  $y_i$

代表重构输出  $Y$  的第  $i$  个元素,  $x_i$  代表卷积自编码网络输入样本  $X$  的第  $i$  个元素。

$L_{SSIM}$  代表代表结构相似性损失, 其计算方式如式 (6) 所示:

$$L_{SSIM} = \frac{(2\mu_Y\mu_X + c_1)(2\sigma_{XY} + c_2)}{(\mu_Y^2 + \mu_X^2 + c_1)(\sigma_Y^2 + \sigma_X^2 + c_2)} \quad (6)$$

式中,  $\mu_X$  和  $\mu_Y$  分别代表输入样本和重构样本的均值, 而  $\sigma_X$  和  $\sigma_Y$  则分别代表输入样本  $X$  和重构样本  $Y$  的标准差,  $\sigma_{XY}$  是  $X$  和  $Y$  的协方差,  $c_1$  和  $c_2$  则是固定的常数, 用于稳定结构相似性损失数值。通过引入结构相似性损失, 可考虑的输入样本的结构信息, 并忽略非结构噪声变化, 增强诊断过程对噪声干扰的鲁棒性。

在检测阶段, 将待检测的牵引变压器声纹样本通过连续小波变换后得到相应的时频特征, 并输入到自编码网络中。由于训练阶段完全基于正常牵引变压器声纹样本, 故卷积自编码网络仅能对正常样本实现低误差重构, 即当输入样本是正常运行情况下的牵引变压器声纹时, 自编码器网络重构误差较低; 而当自编码网络的输入是异常样本时, 其解码器部分无法将异常样本的压缩编码低误差的重构到样本空间, 网络对异常样本的重构误差较大。因此, 通过比较自编码网络重构误差值大小即可实现对异常状态的识别。

传统的卷积自编码器通过堆叠常规卷积层的方式完成构建, 其计算复杂度高, 且对于正常样本的压缩重构能力不足, 易导致对正常声纹样本和故障声纹样本的混淆<sup>[24]</sup>。为提升对牵引变压器声纹样本的检测速率以及检测准确率, 提出 SCF-AE 网络结构, 通过 SCF 卷积替代原有卷积自编码网络中的常规卷积层, 提升自编码网络对牵引变压器声纹特征的学习能力的同时, 避免常规卷积层引入的冗余特征。SCF 卷积由空间融合卷积和通道融合卷积两部分组成。

### 1.2.2 空间融合卷积

空间融合卷积模块的整体结构如图 4 所示, 该模块对输入的特征图采用分离-融合操作。分离过程中, 首先对输入特征图  $F$  进行分组归一化<sup>[25]</sup>, 其计算公式如式 (7) 所示:

$$F_G = \gamma \frac{F - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (7)$$

式中,  $\mu$  和  $\sigma$  分别是特征图  $F$  按照通道维度分组后计算得到的均值和方差,  $F$  将首先通过减去均值除以标准差的操作实现特征的归一化, 而  $\gamma$  和  $\beta$  是可学习参数, 用于帮助模型将特征缩放到合适的分布中, 其参数学习机制和其他神经网络参数类似, 即首先计算损失函数对参数  $\gamma$  和  $\beta$  的梯度, 再沿梯度反方向, 以学习率为步长更新参数, 迭代优化。 $\epsilon$  是一个值很小的常数, 用于避免

计算过程中分母为 0 的情况。分组归一化首先对通道进行分组, 然后对分组后的特征求取均值和标准差。

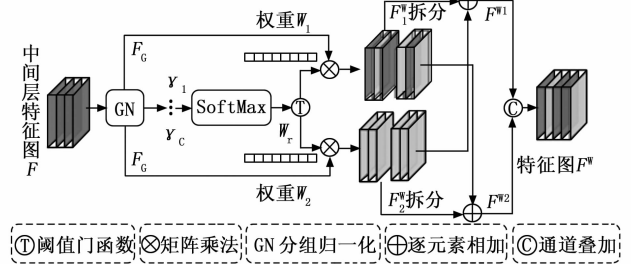


图 4 空间融合卷积模块

在分组归一化过程中, 空间特征的信息量越大, 其数据变化越剧烈,  $\gamma$  就越大。因此, 可通过对参数  $\gamma$  进行 SoftMax 操作计算相应的通道注意力权重  $W_r$ , 并针对权重设置阈值, 其中通道注意力权重  $W_r$  的计算如式 (8) 所示:

$$W_r = \{w_i\} = \frac{\gamma_i}{\sum_{j=1}^C \gamma_j}, i, j = 1, 2, \dots, C \quad (8)$$

设定注意力权重阈值为 0.5, 当  $W_r$  的某个元素小于门限阈值时, 将该元素值置零, 否则保留原始值, 从而得到权重  $W_1$ ; 当  $W_r$  的某个元素大于门限阈值时, 将该元素值置零, 否则保留原始值, 从而得到权重  $W_2$ 。分别使用  $W_1$  和  $W_2$  与归一化特征图  $F_G$  相乘得到两个加权特征图  $F_1^w$  和  $F_2^w$ 。对  $F_1^w$  和  $F_2^w$  分别沿通道方向分为两对, 并将特征按照如图 4 所示的方式进行交叉相加操作, 得到两对特征图  $F^{w1}$  和  $F^{w2}$ , 随后将两对特征图  $F^{w1}$  和  $F^{w2}$  进行逐元素相加, 得到空间融合特征图  $F^w$ 。

### 1.2.3 通道融合卷积

完成空间融合卷积之后, 利用通道融合卷积进一步去除原始特征图中的冗余特征, 其结构如图 5 所示。通道融合卷积的分离操作首先将融合特征图  $F^w$  在通道维度按比例  $\alpha$  拆分为两组, 第一组特征图的通道数为  $\alpha C$ , 而第二组特征图的通道数为  $(1-\alpha)C$ , 在这里  $\alpha$  被设置为 0.5。随后通过  $1 \times 1$  卷积层分别对两组特征图进行特征提取得到特征图  $F^1$  和特征图  $F^2$ 。特征图  $F^1$  将通过全局平均池化和 SoftMax 函数操作, 得到一组特征图注意力权重, 该组注意力权重将与特征图  $F^2$  逐通道相乘, 得到的结果将继续沿通道方向平均分为两组, 两组特征逐元素相加, 得到最终的输出特征图  $F_{out}$ 。

### 1.3 改进自编码网络结构及检测阈值获取

基于 SCF-AE 网络的牵引变压器声纹异常检测模型包含编码器和解码器两部分, 其结构如图 6 所示。

SCF-AE 网络中各层的卷积核大小、通道数等结构参数通过优化得到, 具体的, 对除最后一层外的每个卷积层均设置卷积核大小集合  $[1, 3, 5]$ , 设置输出通道

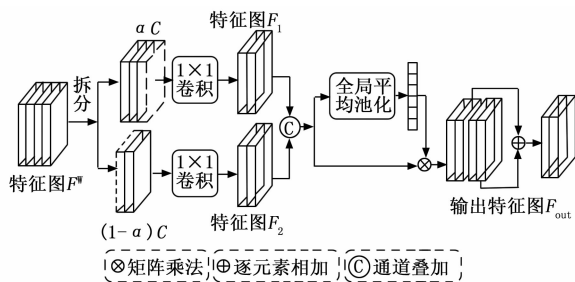


图5 通道融合卷积

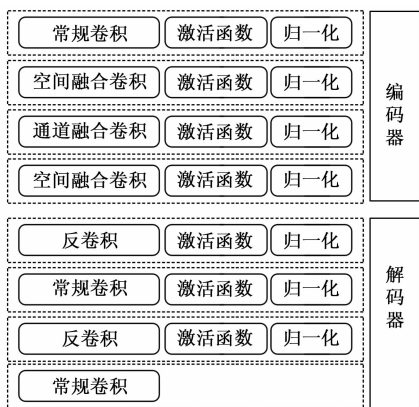


图6 基于 SCF-AE 网络的牵引变压器声纹异常检测模型

数集合 [16, 32, 64, 128], 固定池化层为核大小 2、步长 2 进行网格搜索。搜索过程基于 Optuna 超参数优化库中的多目标网格搜索。搜索过程每种网络参数均会训练 3 轮次, 并基于验证集损失和模型参数数量最小化得到网络各层的参数配置, 设定搜索次数为 40 次。通过优化后, SCF-AE 网络具体结构如表 1, 其编码器和解码器均为四层, 每层网络的具体构成如表 1。除最后一层外, 每个卷积层后都接入 LeakyReLU 激活函数以及批量归一化层, 所有常规卷积层的核大小均为 3, 步长为 1, 边缘填充为 1。编码器下采样通过核大小为 2, 步长为 2 的最大池化完成。

表1 SCF-AE 网络结构

模块	层数	卷积层类型	输出通道数	下采样
编码器	1	常规卷积	16	无
	2	空间融合卷积	32	有
	3	通道融合卷积	32	无
	4	空间融合卷积	32	有
解码器	1	反卷积	32	\
	2	常规卷积	16	\
	3	反卷积	16	\
	4	常规卷积	3	\

完成训练后, 将待检测的变压器原始声纹信号通过 CWT 得到时频谱特征图; 将时频谱特征图输入至 SCF-AE 网络中, 对比重构误差是否超过检测阈值, 若小于

检测阈值则认为属于正常, 若大于检测阈值则认为属于异常, 实现对牵引变压器的声纹异常检测。

在完成最后一轮训练后, 计算所有训练样本 (均为正常样本) 重构误差的均值  $\mu_n$  和标准差  $\sigma_n$ , 并通过式 (9) 所计算网络的异常检测阈值, 该阈值为高斯分布的  $4\sigma_n$  分位数, 在重构误差满足高斯分布的假设下, 正常样本重构误差大于该阈值的概率极低, 可视检测阈值  $T$  为正常样本和异常样本的边界:

$$T = \mu_n + 4\sigma_n \tag{9}$$

式中,  $\mu_n$  为重构误差的均值,  $\sigma_n$  为标准差。

## 2 实验验证及结果分析

### 2.1 实验设置

为验证方法有效性, 通过朔黄铁路发展有限责任公司提供的牵引变压器声纹数据开展试验验证, 根据国标《GB/T 1094.10-2022 电力变压器第 10 部分: 声级测定》所规定位置, 在牵引变压器规定轮廓线上部署声纹传感器进行声纹数据采集<sup>[26]</sup>, 数据包括正常样本及异常样本, 其中异常样本包括老化、过负载及部件松动三类。对获取的数据按长度 1 s 进行切分。正常样本分段后共得到 350 条样本, 异常样本通过切分后共得到 82 条样本。随后, 对所有声纹数据进行 CWT 变换, 得到相应的声纹特征时频特征。对正常样本按照 7 : 3 的比例划分训练集和测试集, 异常样本则全部作为测试集样本。最终的训练集由 245 条由正常牵引变压器声纹特征构成, 而测试集则由 105 条正常样本及 82 条故障样本组成。在实验验证中, 基于 Python3.8 以及 PyTorch2.4.1 深度学习框架搭建自编码神经网络模型, 训练过程中采用 Adam 优化器, 学习率大小为 0.001, 并使用 L2 正则化, 正则化权重设置为 0.000 1, 训练设置为 50 个轮次, 训练批尺寸设置为 16。

### 2.2 牵引变压器声纹异常检测结果

为验证所提出方法的有效性, 在训练集上训练 SCF-AE 网络。在完成训练后, 在测试集上对所提出的 SCF-AE 网络进行测试, 其输出的重构样本如图 7 所示, 可观察到, 对于输入的正常牵引变压器声纹的 CWT 时频特征, SCF-AE 网络可以对其进行精确的重构; 而对于异常牵引变压器声纹样本的时特征, SCF-AE 的重构输出存在明显的异常重构区域, 这些异常重构区域具有明显的不连续和异常值 (表现在时频谱图中即为类似于图像噪声的色块)。由于 SCF-AE 网络对于正常声纹样本的重构误差极小、而对于异常声纹样本的重构误差较大, 因此可通过对比重构误差大小可有效的实现对异常样本的甄别。

在完成训练后, 统计最后一轮训练结束后自编码网

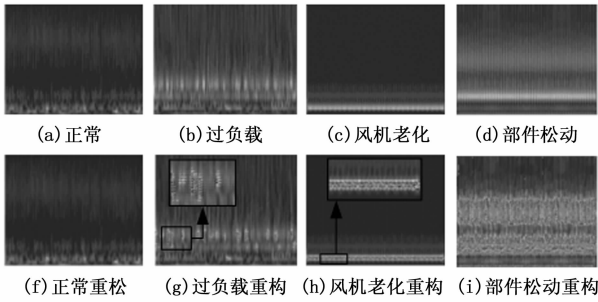


图 7 自编码输出结果和输入的对比

络对训练集样本的重构误差，重构误差均值为 0.026 0，标准差为 0.002 8，因此根据式 (9) 得到的检测阈值为 0.026 8。随后在测试集上验证所提出方法的 AUC-ROC 分数、召回率  $R$  以及精确率  $P$ 。其中 AUC-ROC 分数代表输入正常样本和故障样本时，模型输出重构误差的可区分度，该指标计算和检测阈值无关；而召回率  $R$  和精确率  $P$  则代表通过检测阈值进行分类后，模型分类结果的准确性，这两个指标可说明阈值的合理性。

所提方法在测试集上取得的 AUC-ROC 分数为 100%，说明 SCF-AE 网络对正常样本和异常样本在输出空间上具有较强的区分性；召回率为 100%，说明在该检测阈值下漏检率为 0；精确率为 88.17%，说明存在部分被误诊断为异常的正常样本，但结合 AUC-ROC 分数可知，这种误判是由于阈值计算错误导致，而非网络本身对于正常和异常样本的区分性不足。

### 2.3 对比实验及结果分析

为进一步验证所提出方法性能，与两种无监督方法在相同数据集上进行了对比，对比方法分别为特征记忆库模型<sup>[27]</sup>以及分布外检测神经网络<sup>[28]</sup>，两种方法均无需异常样本训练。其中，特征记忆库模型检测首先通过预训练神经网络提取声纹特征，然后将特征存储为特征记忆库，并通过检索比对的方式完成检测；分布外检测神经网络则是训练神经网络将正常样本映射到一个超球面，并优化超球面的体积。对比结果如表 2。

表 2 与其他无监督学习方法的对比

模型	AUC-ROC/%	精确率/%	召回率/%	推理速度/ms
SCF-AE	100	88.17	100	13.3
特征记忆库	99.80	94.25	100	42.9
One-Class	96.72	100	87.80	14.4

可以观察到，对比其他方法，SCF-AE 在 AUC-ROC 分数、召回率及推理速度等性能指标上具有优势，说明 SCF-AE 在对牵引变压器正常和异常样本的区分性更强，图 8 所示为 SCF-AE 对正常样本和异常样本的输出重构误差值散点图，图中红线标记出了通过式 (6)

得到的检测阈值界限。

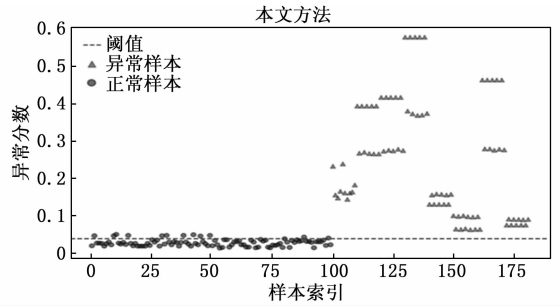


图 8 输出值散点图及阈值界限

从图 8 中可看到，导致 SCF-AE 在精确率上不足的原因是设置的阈值过小，将部分正常的牵引变压器声纹样本误判为异常，因此，通过选取更合理的阈值，所提方法具备实现精确率和召回率均衡的可能。特征记忆库模型和分布外检测网络的预测散点值和检测阈值则如分别如图 9 及 10 所示，从中可以看到，尽管两种方法具有更好的精确率，但两种方法本身会对部分异常样本和正常样本产生混淆，因此，通过调整阈值无法实现更进一步的性能提升，注意到由于方法不同，因此对比方法和 SCF-AE 网络输出的异常分数数值上存在较大差异。

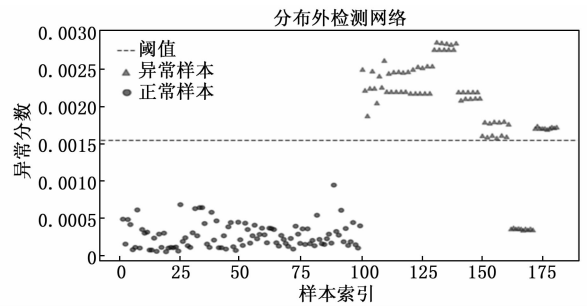


图 9 分布外检测网络输出值散点图及阈值界限

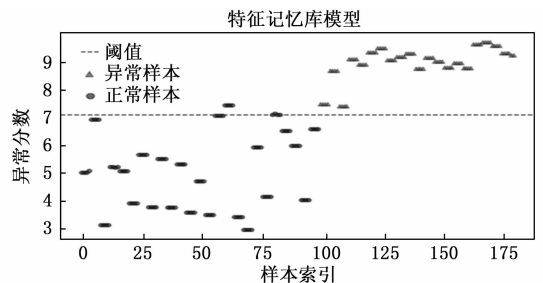


图 10 特征记忆库模型输出值散点图及阈值界限

### 2.4 消融实验及结果分析

SCF-AE 网络分别对声纹特征提取方法和自编码网络结构进行了改进，为验证所作改进的有效性，进行相应消融实验。不同消融模型的性能如表 3 所示，其中 AE 为传统卷积自编码网络，STFT 代表采用短时傅里

叶变换特征提取，CWT 则是通过连续小波变换提取声纹特征。

表 3 不同消融模型性能

	AE+STFT	AE+CWT	SCF-AE+STFT	SCF-AE+CWT
AUC-ROC/%	95.99	95.94	95.96	100
Precision/%	92.16	80.43	90.91	88.17
Recall/%	57.32	90.24	48.78	100
参数数量	33 603	33 603	18 483	18 483

可以观察到，CWT 变换特征提取对诊断性能的影响最大，SCF-AE 网络的诊断性能相较于原始的自编码网络得到提升，且 SCF-AE 网络参数数量更低。

随后，对两种模型在 STFT 变换上的部分预测结果进行进一步分析，如图 11 所示为两种模型的部分预测结果，从图中结果中可以看到，对于正常、绕阻老化、过负载以及部件松动的 STFT 变换特征谱图，传统自编码网络都难以很好的进行压缩重构，因此无法实现对异常和正常样本的区分；SCF-AE 网络具有更强的特征学习能力，而 STFT 变换谱图时频分辨率不足，信息量少，因此无论正常还是异常 SCF-AE 网络都可很好重构，无法实现对异常和正常样本的区分。

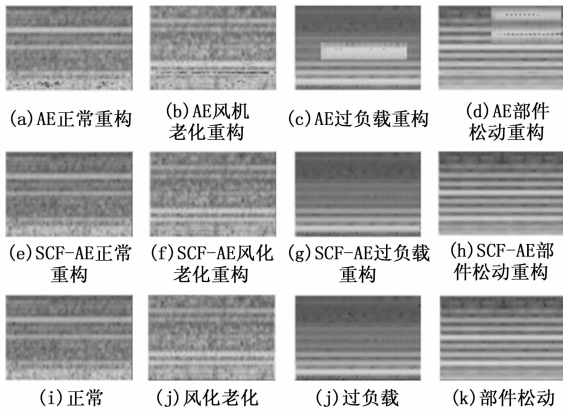
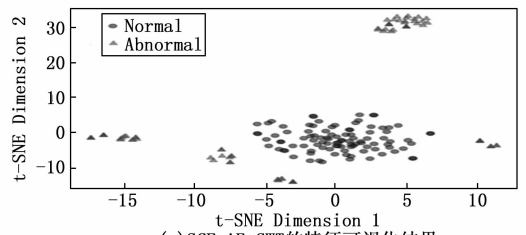


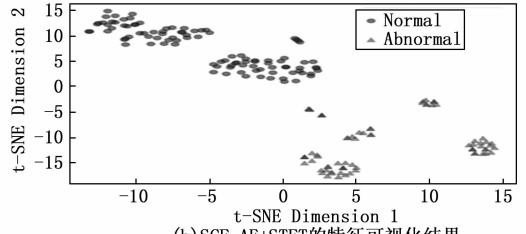
图 11 SCF-AE 与 AE 的重构输出对比

### 2.5 消融模型特征可视化分析

为进一步分析不同消融模型的性能差异，探究不同消融模型性能差异的内在因素，对采用不同特征提取方法、不同模型组合的消融模型编码器输出特征进行了 t-SNE 特征降维可视化操作，从而直观的分析正常和异常样本在不同消融模型特征空间上的差异性。对所提出的 SCF-AE 网络，分别采用 STFT 和 CWT 进行特征提取，所得到的特征可视化结果如图 12 所示，可明显观察到采用 CWT 时正常样本在 SCF-AE 特征空间上的分布更紧凑，便于模型的判别；而采用 STFT 时，正常样本分布离散，正常样本分布的边界求取存在困难。



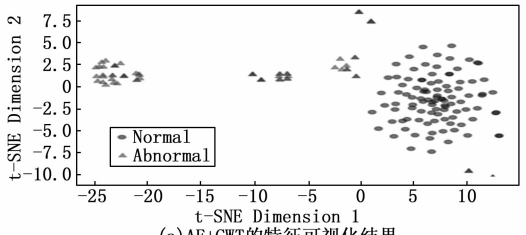
(a) SCF-AE+CWT 的特征可视化结果



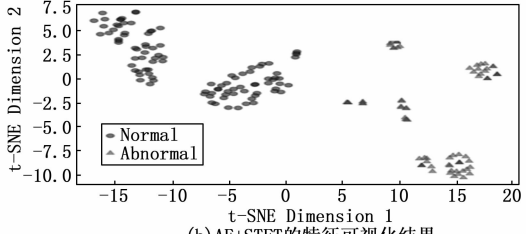
(b) SCF-AE+STFT 的特征可视化结果

图 12 SCF-AE 编码层 t-SNE 特征可视化

而对于简单的 AE 网络，同样分别采用 STFT 和 CWT 进行特征提取，所得到的特征可视化结果如图 13 所示，可明显观察到采用 CWT 时正常样本在 AE 特征空间上的分布相比 SCF-AE 更松散，但明显存在一个聚类中心；而采用 STFT 时，正常样本分散，不同正常样本距离更远，导致判别边界的求取存在困难。



(a) AE+CWT 的特征可视化结果



(b) AE+STFT 的特征可视化结果

图 13 AE 编码层 t-SNE 特征可视化

### 3 结束语

为解决牵引变压器故障声纹样本稀缺问题，提出了结合 CWT 变换及 SCF-AE 网络的牵引变压器异常声纹检测方法。首先，利用 CWT 变换提取高分辨率的声纹时频特征；随后提出引入空间通道融合卷积模块改进的 SCF-AE 网络，提升原始自编码器的特征学习能力，并降低计算复杂度；经测试，提出的 SCF-AE 网络在数据集上取得了 100% 的 AUC-ROC 分数、100% 召回率以及 88.17% 的精确率，可实现准确的异常声纹检测，所提方法在训练阶段无需故障声纹数据且具有较强的判别

能力,避免了因样本稀缺导致的神经网络性能退化情况,在后续研究中,将考虑采用模糊辨识等更先进方法,深入考虑如何在故障样本稀缺条件获取具有鲁棒诊断能力的阈值边界。

#### 参考文献:

- [1] 马 帅,刘秀波,任松斌,等.朔黄重载铁路曲线钢轨侧磨预测研究 [J]. 机械工程学报, 2021, 57 (18): 118 - 125.
- [2] 张志刚.朔黄铁路牵引变压器在线监测技术研究及应用 [D]. 成都:西南交通大学, 2013.
- [3] 李林达,李正绪,孙实源,等.电力变压器短路累积效应研究综述 [J]. 变压器, 2017, 54 (2): 24 - 31.
- [4] ELMOUDI A, LEHTONEN M, NORDMAN H. Effect of harmonics on transformers loss of life [C] //Conference Record of the IEEE International Symposium on Electrical Insulation, IEEE, 2006.
- [5] 江俊飞.高速铁路大型变压器绕组频率响应建模及故障诊断研究 [D]. 成都:西南交通大学, 2019.
- [6] 刘志坚,何 蔚,刘 航,等.基于格拉姆角场变换和深度压缩模型的变压器故障识别方法 [J]. 电网技术, 2023, 47 (4): 1478 - 1490.
- [7] 李 刚,米琛浩,郑顾平,等.隐式半马尔科夫模型下的变压器故障诊断方法 [J]. 计算机测量与控制, 2019, 27 (5): 30 - 34.
- [8] 温阳东,宋 阳,王颖鑫,等.基于模糊神经网络的电力变压器故障诊断 [J]. 计算机测量与控制, 2013, 21 (1): 39 - 41.
- [9] 汲胜昌,张 凡,师榆航,等.基于振动信号的电力变压器机械状态诊断方法研究综述 [J]. 高电压技术, 2020, 46 (1): 257 - 272.
- [10] 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, 2023.
- [11] 包艳艳,杨广泽,陈 伟,等.基于 SBSS 与 CNN 的 750 kV 变压器与尖板放电混叠信号的声纹识别 [J]. 西南交通大学学报, 2023: 1 - 11.
- [12] YU Z, WEI Y J, NIU B, et al. Automatic condition monitoring and fault diagnosis system for power transformers based on voiceprint recognition [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1 - 11.
- [13] 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.
- [14] 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 - 423.
- [15] 马宏忠,李 楠,杨启帆,等.基于多特征声纹图谱的变压器绕组松动在线故障诊断方法 [J]. 电机与控制学报, 2023, 27 (5): 76 - 87.
- [16] 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, 2023: 1 - 13.
- [17] 仲林林,胡 霞,刘柯好.基于改进生成对抗网络的无人机电力杆塔巡检图像异常检测 [J]. 电工技术学报, 2022, 37 (9): 2230 - 2240.
- [18] 万可力,马宏忠,崔佳嘉,等.基于 Mel-GADF 与 ConvNeXt-T 的变压器铁心松动故障诊断方法 [J]. 电力自动化设备, 2024, 44 (3): 217 - 224.
- [19] 崔佳嘉,马宏忠.基于改进 MFCC 和 3D-CNN 的变压器铁心松动故障声纹识别模型 [J]. 电机与控制学报, 2022, 26 (12): 150 - 160.
- [20] LI J F, WEN Y, HE L H. SCConv: spatial and channel reconstruction convolution for feature redundancy [C] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 6153 - 6162.
- [21] 王嘉宁.基于无监督学习的电力变压器声纹异常诊断方法研究 [D]. 北京:华北电力大学, 2023.
- [22] LI P Z, YAN P, LI J I. A comprehensive survey on design and application of autoencoder in deep learning [J]. Applied Soft Computing, 2023, 138: 110176.
- [23] 张森彬,王丰华,金玉琪,等.基于图像分割及小波脊线的变压器绕组状态检测 [J]. 电工技术学报, 2025, 40 (2): 640 - 652.
- [24] HUO Y, CHENG X, LIN S, et al. Memory-Augmented autoencoder with adaptive reconstruction and sample attribution mining for hyperspectral anomaly detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1 - 18.
- [25] HUANG L, QIN J, ZHOU Y, et al. Normalization techniques in training DNNs: methodology, analysis and application [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45 (8): 10173 - 10196.
- [26] GB/T 1094.10-2022, 电力变压器第 10 部分: 声级测定 [S]. 2022.
- [27] ROTH K, PEMULA L, ZEPEDA J, et al. Towards total recall in industrial anomaly detection [C] //2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 14298 - 14308.
- [28] RUFF L, VANDERMEULEN R, GOERNITZ N, et al. Deep One-Class classification [C] //Proceedings of the 35th International Conference on Machine Learning. PMLR, 2018: 4393 - 4402.