

DOI: 10.12158/j.2096-3203.2023.02.023

基于改进半监督阶梯网络的有载分接开关故障诊断方法

郑尚直¹, 仲林林¹, 王同磊², 高丙团¹

(1. 东南大学电气工程学院, 江苏 南京 210096;

2. 国网江苏省电力有限公司电力科学研究院, 江苏 南京 211103)

摘要:有载分接开关(on-load tap changer, OLTC)是变压器实现有载调压的关键部件,具有复杂的机电结构。基于振动信号的OLTC机械故障诊断目前存在样本标记难度高而难以有效训练的难题。文中提出一种基于贝叶斯优化和卷积算子改进的阶梯网络(Bayesian optimization-convolutional ladder networks, BO-ConvLN),可在少标签情况下提高OLTC机械故障诊断的精度。首先,引入阶梯网络对振动信号进行半监督学习,充分利用大量无标签样本的潜在信息对特征提取过程进行指导,提升少标签情况下阶梯网络的诊断性能。然后,用卷积算子替代阶梯网络中的全连接形式的线性变换,增强阶梯网络对非平稳振动信号的特征提取能力。在此基础上,借助贝叶斯优化对阶梯网络高维超参数进行寻优,在有限时间成本下显著提升了模型的诊断精度。实验结果显示,在仅有40个标签的情况下,对传动卡涩、芯子润滑不足、顶盖松动3类故障的诊断精度达91.67%,证明了BO-ConvLN在故障诊断中的有效性。

关键词:有载分接开关(OLTC);故障诊断;振动信号;阶梯网络;半监督学习;贝叶斯优化

中图分类号:TM407

文献标志码:A

文章编号:2096-3203(2023)02-0197-09

0 引言

有载分接开关(on-load tap changer, OLTC)是变压器中唯一可以动作的部件,具有复杂的机械结构和电气特性^[1-2]。据统计,因OLTC故障引起的变压器事故占变压器总事故的18.5%,其中紧固件松脱、异常机械磨损、异常转矩与传动卡涩等机械故障占比70%~90%,异常过电压、绝缘油劣化等电气故障占比10%~30%^[3],且许多电气故障也都由机械故障发展而来。因此,对其机械性能进行有效评估对提高OLTC乃至变压器的运行可靠性具有重要意义。

振动信号是OLTC状态估计的主要测量量^[4-5]。OLTC振动信号可通过在线监测、离线检测和故障模拟实验获取,在线监测获取的样本数量多,但很难进行有效标注;离线检测和故障模拟实验能有效标注样本,但可获取的样本数量少,因此OLTC故障诊断往往面临样本标签数量少的问题^[6]。而传统机器学习方法在标签数量较少的情况下也具有较好的分类性能,常被应用于OLTC的故障诊断^[7]。由于OLTC振动信号是具有强非线性的非平稳信号,常规的时域或频域分析方法难以有效提取混叠的信息。文献[8-10]利用希尔伯特-黄变换(Hibert-Huang transform, HHT)或经验模态分解(empirical mode decomposition, EMD)对OLTC振动信号进行处

理,提取了包络谱熵、低频包络、功率矩阵相似度等特征,实现了对弹簧松动、触头磨损、绝缘板松动等故障的诊断。文献[11-12]基于相空间重构法,提取了相空间轨迹矩阵奇异值和相点空间分布系数2类特征,实现了对储能弹簧力下降、软连接松动等故障的诊断。基于传统机器学习的OLTC故障诊断方法虽然在一定程度上缓解了对标签数量的需求,但仍要基于专家经验对每类故障的特征进行人工设计,且特征优劣将直接决定诊断效果。

深度学习方法具有强大的非线性映射能力,能根据样本数据的内在规律,自动提取与分类任务最相关的特征。文献[13]提取振动信号的增强递归图作为卷积神经网络(convolutional neural network, CNN)的输入,实现了对4种典型故障的诊断。文献[14]将振动信号转换为时间-加速度图像,通过CNN提取图像特征,再将所提特征与人工特征融合,用于故障诊断。上述文献初步探索了采用深度神经网络自动提取OLTC振动信号特征,但由于标签数量较少,通过监督学习方法训练深度神经网络的诊断精度受到了限制。

为有效利用大量无标签样本的潜在信息,提升基于深度学习的OLTC故障诊断模型的诊断精度,文中提出了一种基于贝叶斯优化和卷积算子改进的阶梯网络(Bayesian optimization-convolutional ladder networks, BO-ConvLN)模型。引入阶梯网络对OLTC振动信号进行半监督学习,充分利用大量无标签样本的潜在信息对特征提取过程进行指导,可

收稿日期:2022-09-24;修回日期:2022-12-01

基金项目:国家自然科学基金资助项目(92066106)

提升少标签情况下阶梯网络的诊断性能;对阶梯网络进行卷积化改造,可增强阶梯网络对强非线性、非平稳振动信号的特征提取能力;借助贝叶斯优化对阶梯网络高维超参数进行寻优,在有限时间成本下显著提升了模型的诊断精度。

1 卷积半监督阶梯网络

1.1 半监督学习原理

半监督学习^[15-16]是一类结合监督学习和无监督学习的机器学习范式,能同时利用有标签和无标签样本进行训练。半监督学习的原理可用如图1所示的二分类问题进行阐述,其中黄色和绿色小球代表类别不同的有标签样本,灰色小球代表无标签样本。监督学习方法只能利用2个有标签样本训练二分类器,大量无标签样本的潜在信息未被有效利用,此时训练得到的二分类器分类边界如虚线所示,分类器的预测精度非常低。因此在监督学习的基础上,半监督学习引入无监督学习来挖掘2类样本的分布情况,进一步优化二分类器的分类边界,如实线所示,大大提高了分类器的预测精度。

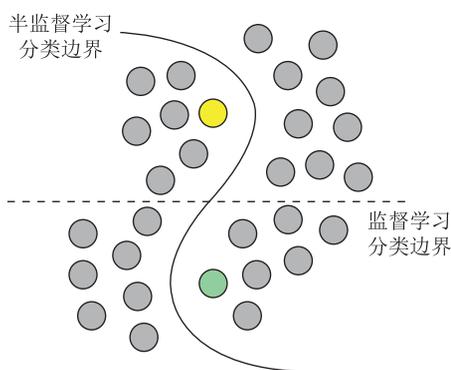


图1 半监督学习原理

Fig.1 Principle of semi-supervised learning

1.2 半监督阶梯网络结构

阶梯网络是一种用于半监督学习的深度学习架构。3层阶梯网络的结构如图2所示,包含2个编码器和1个解码器,其中一个编码器与解码器之间存在横向连接,整个网络结构形似“梯子”^[17]。其中, v 为OLTC振动信号; $z_l, \tilde{z}_l, \hat{z}_l$ 分别为第 l 层的无噪编码器输出、含噪编码器输出和解码器输出; h_l, \tilde{h}_l 分别为 z_l, \tilde{z}_l 经过仿射变换和非线性激活后的结果; $C_{d,l}$ 为第 l 层的均方误差(mean square error, MSE); \tilde{y}, y 分别为含噪编码器和无噪编码器的输出。

阶梯网络的监督学习过程由图2左侧的含噪编码器实现,每层含噪编码器包含线性变换、批归一

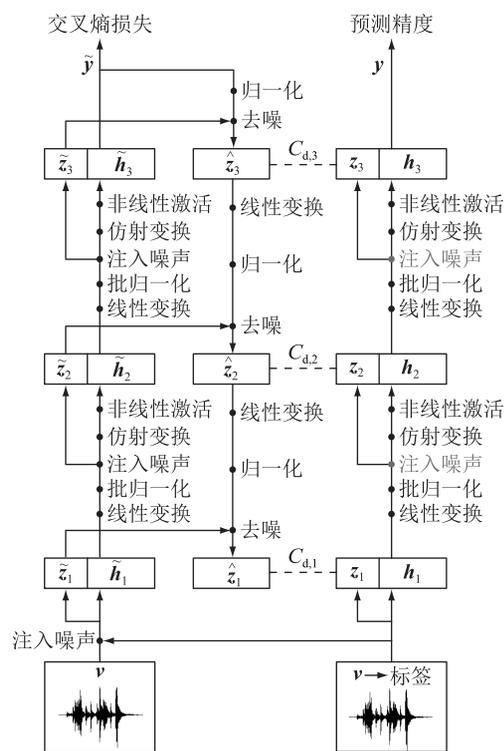


图2 3层阶梯网络的结构

Fig.2 Structure of three-layer ladder networks

化、注入噪声、仿射变换、非线性激活5种操作。设有标签样本为 $\{v_n, e_n | 1 \leq n \leq N\}$,其中 N 为有标签样本数量; v_n 为样本数据; e_n 为对应的标签。在 v 中注入高斯噪声,并将其作为含噪编码器的输入 \tilde{h}_1 。线性变换操作将下层编码器的输出 \tilde{h}_{l-1} 通过权重矩阵 W_l 线性映射到本层编码器激活前的输出 \tilde{z}_l^{pre} 。批归一化和仿射变换操作使得 \tilde{z}_l^{pre} 的均值为0,标准差为1,从而避免神经网络“梯度消失”的问题^[18]。注入噪声操作在 \tilde{z}_l^{pre} 中注入高斯噪声得到含噪特征 \tilde{z}_l ,注入噪声的目的是借助噪声对训练过程进行正则化。非线性激活操作通常采用ReLU激活函数,但对于编码器的最高层则采用Softmax激活函数,激活后得到本层编码器的输出 \tilde{h}_l 。将最高层编码器的输出 \tilde{h}_l 用于预测样本的标签 \tilde{y} ,监督学习过程的损失函数 C_c 即为预测标签 \tilde{y} 和真实标签 e_n 之间的交叉熵^[19]:

$$C_c = -\frac{1}{N} \sum_{n=1}^N \ln P(\tilde{y} = e_n | v_n) \quad (1)$$

式中: $P(\cdot)$ 为概率函数。

阶梯网络的无监督学习过程由图2中间的解码器和右侧的无噪编码器实现。无噪编码器与含噪编码器结构类似,两者共享一个权重矩阵 W_l ,唯一不同的是无噪编码器没有注入噪声操作。每层解码器包含线性变换、归一化、去噪3种操作。设全体

样本为 $\{\mathbf{v}_n \mid 1 \leq n \leq M\}$, 其中 M 为全体样本数量。对于无噪编码器, 直接将 \mathbf{v} 作为其输入 \mathbf{h}_1 ; 对于解码器, 将 $\tilde{\mathbf{h}}_L$ 进行归一化和去噪后作为其输入 $\hat{\mathbf{z}}_L$ 。线性变换和归一化操作将上层解码器的输出 $\hat{\mathbf{z}}_{l+1}$ 通过权重矩阵 \mathbf{V}_l 线性映射到 \mathbf{u}_l 。去噪操作通过横向连接, 并结合 \mathbf{u}_l 和 $\tilde{\mathbf{z}}_l$ 的信息构造出本层解码器的输出 $\hat{\mathbf{z}}_l$, 试图通过 $\hat{\mathbf{z}}_l$ 来估计无噪编码器的输出 \mathbf{z}_l 。无监督学习过程的损失函数 C_d 为解码器输出 $\hat{\mathbf{z}}_l$ 与无噪编码器输出 \mathbf{z}_l 之间的 MSE:

$$C_d = \sum_{l=1}^L \lambda_l C_{d,l} = \sum_{l=1}^L \frac{\lambda_l}{M m_l} \sum_{j=1}^{N_T} \|\mathbf{z}_l(j) - \hat{\mathbf{z}}_l(j)\|^2 \quad (2)$$

式中: L 为神经网络的层数; λ_l 为第 l 层神经网络在无监督损失中的权重; m_l 为第 l 层神经网络的神经元个数; $\mathbf{z}_l(j)$ 、 $\hat{\mathbf{z}}_l(j)$ 分别为 \mathbf{z}_l 、 $\hat{\mathbf{z}}_l$ 的第 j 个分量; N_T 为神经元总数。

则半监督阶梯网络的损失函数 C 为:

$$C = C_c + C_d \quad (3)$$

1.3 半监督阶梯网络的卷积化改造

阶梯网络中线性变换采用全连接形式, 权重矩阵 \mathbf{W}_l 和 \mathbf{V}_l 的参数数量取决于相邻 2 层神经网络的神经元个数。设编码器第 $l-1$ 层、 l 层的神经元个数分别为 m_{l-1} 、 m_l , 则编码器第 l 层权重矩阵 \mathbf{W}_l 的参数数量为 $m_{l-1} m_l$, 且 m_0 为样本的数据长度。由此可见, 阶梯网络权重矩阵的参数数量正比于样本数据长度的平方。文中将 OLTC 切换开关动作时所对应的振动信号作为阶梯网络的输入, 振动信号 \mathbf{v} 的采样时长为 0.3 s, 采样率为 100 kHz, 数据长度为 30 000。由此估计, 全连接阶梯网络的参数数量十分庞大, 将大大增加训练模型所需的计算开销和阶梯网络的过拟合风险^[20]。同时, 全连接阶梯网络中大量参数都被用于提取振动信号的全局特征, 但 OLTC 振动信号具有非平稳、非周期的特点, 振动信号的幅值频率随时间快速变化, 不同时间段内振动信号的关联性较弱, 因此全连接阶梯网络对 OLTC 振动信号的特征提取效率很低。为减少阶梯网络在样本数据长度较长时的参数数量, 提升阶梯网络对非平稳、非周期振动信号的特征提取效率, 文中借鉴卷积神经网络中参数共享的思想, 用卷积算子替代阶梯网络中的全连接形式的线性变换。

对于编码器, 线性变换过程被替换为如式(4)、式(5)所示的形式^[21]。

$$\mathbf{z}_l^{\text{pre}} = \text{Conv}_l(\mathbf{h}_l) \quad (4)$$

$$\tilde{\mathbf{z}}_l^{\text{pre}} = \text{Conv}_l(\tilde{\mathbf{h}}_{l-1}) \quad (5)$$

式中: $\text{Conv}_l(\cdot)$ 为第 l 层的卷积运算。

对于解码器, 线性变换过程被替换为如式(6)所示的形式。

$$\begin{cases} \mathbf{u}_l = N_B(\tilde{\mathbf{h}}_l) & l = L \\ \mathbf{u}_l = N_B(\text{Conv}_{l+1}(\tilde{\mathbf{z}}_{l+1})) & l < L \end{cases} \quad (6)$$

式中: $N_B(\cdot)$ 为归一化操作。

2 阶梯网络超参数的贝叶斯优化方法

2.1 贝叶斯优化理论

贝叶斯优化是一种基于高斯过程的序贯优化方法, 可在有限的评估开销下得到目标函数的近似最优解^[22-23]。贝叶斯优化问题可描述为在参数空间中寻找一组参数 \mathbf{x} 使得目标函数 $f(\mathbf{x})$ 尽可能大。贝叶斯优化通过高斯过程建立 $f(\mathbf{x})$ 的概率代理模型, 并根据采集函数确定新的评估点 \mathbf{x}_* , 计算 \mathbf{x}_* 对应的目标函数值 f_* , 再将评估后得到的 (\mathbf{x}_*, f_*) 用来更新概率代理模型, 如此循环。

对于由参数 \mathbf{X} 和对应目标函数 y 组成的样本 (\mathbf{X}, y) 训练得到的高斯过程, 可以用贝叶斯公式推导评估点 (\mathbf{x}_*, f_*) 的后验分布^[24]:

$$f_* \mid \mathbf{X}, y, \mathbf{x}_* \sim N(u_*, \sigma_*^2) \quad (7)$$

式中: u_* 、 σ_*^2 分别为 f_* 所服从的高斯过程的均值和方差。

在贝叶斯优化中, 下一个评估点由采集函数确定。期望提升 (expected improvement, EI) 是贝叶斯优化中最常用的采集函数之一^[25-26]。设前 k 次评估中目标函数的最大值为 f_k^{\max} , 下一次评估得到的目标函数值为 f_* , 那么下一次评估对目标函数带来的提升为 $\max(f_* - f_k^{\max}, 0)$ 。EI 采集函数 $a_l(\mathbf{x}_*)$ 可表示为:

$$a_l(\mathbf{x}_*) = E(\max(f_* - f_k^{\max}, 0)) \quad (8)$$

式中: $E(\cdot)$ 为期望函数。

选择采集函数最大值所对应的参数作为下一次评估的参数, 记为 \mathbf{x}_{l+1} 。

$$\mathbf{x}_{l+1} = \text{argmax } a_l(\mathbf{x}_*) \quad (9)$$

2.2 阶梯网络超参数的贝叶斯优化流程

相较于常见的神经网络结构, 阶梯网络除了包含神经网络的层数 L 、每层神经网络的神经元个数 m_1, \dots, m_L 等基本的超参数外, 还包含编码器中注入高斯噪声的幅值 A_{noise} 、每层解码器重构损失的权重 $\lambda_1, \dots, \lambda_L$ 这 2 类特殊超参数。而这 2 类超参数对阶梯网络的性能影响较大, 因此有必要对其单独优化。随着网络层数 L 的增加, $\mathbf{x} = [A_{\text{noise}} \ \lambda_1 \ \dots \ \lambda_L]$

的维数和目标函数 $f(\mathbf{x})$ ($f(\mathbf{x})$ 为超参数为 \mathbf{x} 的阶梯网络在验证集上的测试精度) 的评估时间也会增加, 这对优化算法提出了很高的要求。在超参数向量维度不超过 20 的情况下, 贝叶斯优化算法能在较少的评估次数下获得相对较好的优化效果。文中采用贝叶斯优化算法对 $\mathbf{x} = [A_{\text{noise}} \lambda_1 \cdots \lambda_L]$ 进行优化的流程如图 3 所示, 其中 $\text{GP}(\cdot)$ 为高斯过程; $m(\mathbf{x})$ 为高斯过程的期望; $k(\mathbf{x}, \mathbf{x}^T)$ 为高斯过程的协方差。首先, 基于高斯过程建立目标函数的概率代理模型, 并在优化空间中采样 k 个初始评估点进行初始评估, 为使初始样本更均匀地分布在超参数空间中, 文中采用 Sobol 序列^[27] 采样初始评估点。其次, 根据评估结果 D 更新概率代理模型的后验分布, 并基于采集函数确定下一个评估点 \mathbf{x}_{i+1} 。然后, 将最新的评估结果 ($\mathbf{x}_{i+1}, f(\mathbf{x}_{i+1})$) 加入到 D 中, 并跳转到更新后验分布步骤进行循环。最后, 输出使目标函数最大的评估点作为优化结果。

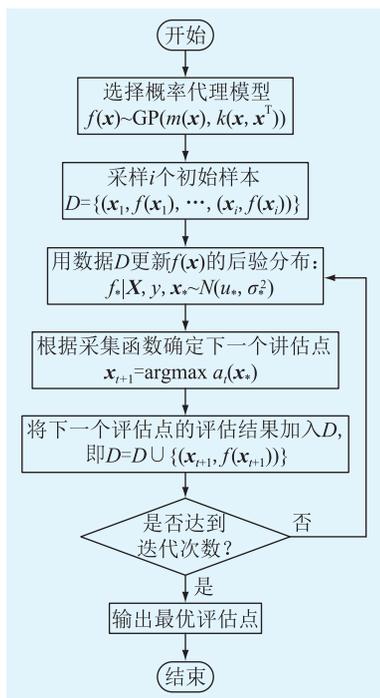


图 3 阶梯网络超参数的贝叶斯优化流程

Fig.3 Flow chart of hyperparameter optimization for ladder network based on Bayesian optimization

3 实验与分析

为与当前先进的半监督学习方法进行对比, 文中首先在智能维护系统 (intelligent maintenance systems, IMS) 数据集^[28] 上进行相关实验研究, 然后为测试 BO-ConvLN 在 OLTC 故障诊断上的性能, 对国内某换流站采集的 OLTC 数据进行实验。

3.1 IMS 数据集

IMS 数据集是工业设备故障诊断领域中应用最

广泛的数据集, 文中使用 IMS 数据集测试阶梯网络和贝叶斯优化方法在振动信号故障诊断中的效果。采集 IMS 数据集的实验平台如图 4 所示, 在该实验平台的电机转轴上安装 4 个双列轴承, 并向每个轴承施加 6 000 lb 的径向载荷。实验共进行 3 次, 每次实验电机转动超过 100 000 000 圈。初始阶段轴承状态良好, 随着转动圈数超过设计极限, 部分轴承逐渐发生故障, 此时电机还会继续转动, 直至达到实验设定的转动圈数。实验中通过加速度传感器采集轴承径向的振动信号, 采样频率为 20 kHz, 并将每 10 min 采集到的振动信号保存到一个文件中。

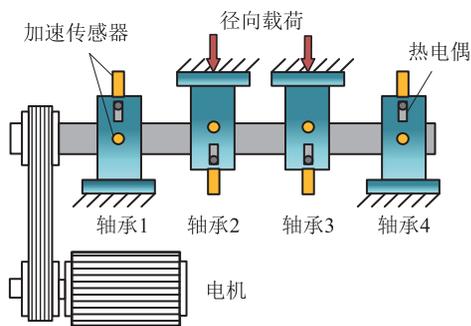


图 4 IMS 数据集实验平台

Fig.4 Experiment setup of IMS dataset

文中选取每类故障发生时对应的 210 个文件进行实验。正常样本选自第 1 次实验中轴承 3 的第 1—210 号文件; 内圈故障样本选自第 1 次实验中轴承 3 的第 1 832—2 041 号文件; 外圈故障样本选自第 2 次实验中轴承 1 的第 352—561 号文件; 滚珠故障样本选自第 1 次实验中轴承 4 的第 1 446—1 655 号文件。将每个文件中的振动信号均等切分为 20 个振动信号片段, 4 类振动信号可切分出 16 800 个样本, 每个样本包含 1 024 个采样点。其中 16 000 个样本用于训练, 800 个样本用于测试。

文中对 IMS 数据集共进行 9 组实验, 每组实验中抽取部分 (分别为 40 个、100 个、200 个、400 个、800 个、1 000 个、2 000 个、4 000 个、8 000 个) 训练样本进行标记, 用于监督学习; 所有的有标签和无标签样本都会用于无监督学习。由于 IMS 数据集中每个样本仅包含 1 024 个采样点, 且振动信号比较平稳, 文中采用未经卷积算子改进的贝叶斯优化阶梯网络 (Bayesian optimization-ladder networks, BO-LN) 和阶梯网络进行实验。每组实验均对 BO-LN 和阶梯网络进行 10 轮训练和测试, 取 10 轮测试精度的平均值作为测试结果。2 类模型的参数如下: 阶梯网络的编码器、解码器采用对称的 7 层全连接结构, 每层神经元个数为 1 024、1 500、500、

250、250、150、4, $\lambda_1 \sim \lambda_7$ 分别为 1 000、10、0.1、0.1、0.1、0.1、0.1, A_{noise} 为 0.6 ($\lambda_1 \sim \lambda_7$ 和 A_{noise} 通过手动调参得到);训练时采用 Adam 优化算法,学习率为 0.02(从第 15 代开始衰减),训练次数为 1 500。BO-LN 的网络结构与阶梯网络相同, $\lambda_1 \sim \lambda_7$ 优化范围为 $0.1\lambda \sim 10\lambda$ (λ 为超参数 λ_L 的基准值), A_{noise} 优化范围为 0.1~1.0;训练方法与阶梯网络相同,贝叶斯优化时采用 20 次 Sobol 采样和 10 次 EI 优化。将 BO-LN 和阶梯网络的测试结果与文献[29]中的变分自编码器(variational autoencoder, VAE) M1、VAE M2、主成分分析(principal component analysis, PCA)+支持向量机(support vector machine, SVM)、自编码器、CNN 进行比较,结果如图 5 所示。由于标签数成倍增加,为避免少标签情况下实验数据过于靠近,图 5 横坐标采用对数坐标。

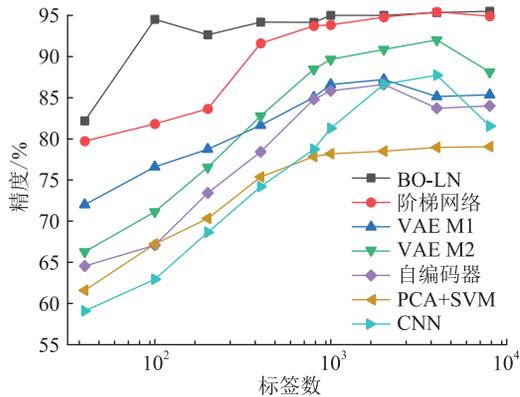


图 5 IMS 数据集实验结果

Fig.5 Experiment results of IMS dataset

由图 5 可知,半监督学习模型(BO-LN、阶梯网络、VAE M1、VAE M2、自编码器)相较监督学习模型(PCA+SVM、CNN)具有更高的精度,其中阶梯网络的平均诊断精度比 CNN 高 14.29%,说明结合无标签样本对监督学习进行正则化将有效提高模型的诊断精度。阶梯网络在 IMS 数据集上的诊断效果要优于其他半监督学习模型(VAE M1、VAE M2、自编码器),通过贝叶斯优化算法对阶梯网络的超参数进行优化能在有标签样本较少的情况下显著提升网络性能,而 BO-LN 的平均诊断精度比阶梯网络高 3.2%。

为检验贝叶斯优化的效果,文中在 IMS 数据集上进行了基于随机搜索^[30]的超参数优化实验。随机搜索在优化空间中对超参数进行有限次随机采样,并选取使得目标函数最大的超参数作为优化结果。图 6 为贝叶斯优化(20 次 Sobol 采样和 10 次 EI 优化)与随机搜索(30 次随机采样)在验证集上测试精度的对比。由图 6 可知,在相同迭代次数下,相

较随机搜索,贝叶斯优化通常能得到在验证集上表现更好的超参数。

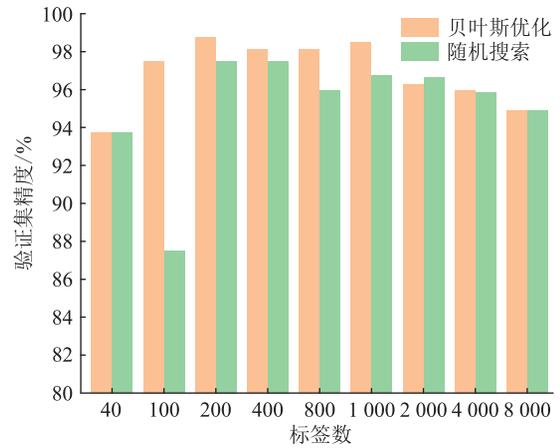


图 6 贝叶斯优化与随机搜索对比

Fig.6 Comparison between Bayesian optimization and random search

3.2 OLTC 数据集

在国内某换流站采集了多个 VRG 型 OLTC 的振动信号,其切换开关如图 7(a)所示,加速度传感器的安装位置如图 7(b)所示。数据采集装置采用 CompactDAQ 系统,该系统包含 1 个 NI 9215 机箱和 2 块 NI 9250 数据采集卡。振动信号的采样频率为 100 kHz,采样时间为 7 s。数据集中包含 4 类振动信号,分别为正常、传动卡涩、芯子润滑不足、顶盖松动,每类包含 60 个样本,共 240 个样本。其中 200 个样本用于训练,40 个样本用于测试。训练时,提取切换开关动作时刻对应的振动信号主波形作为模型的输入,部分振动信号主波形如图 8 所示。

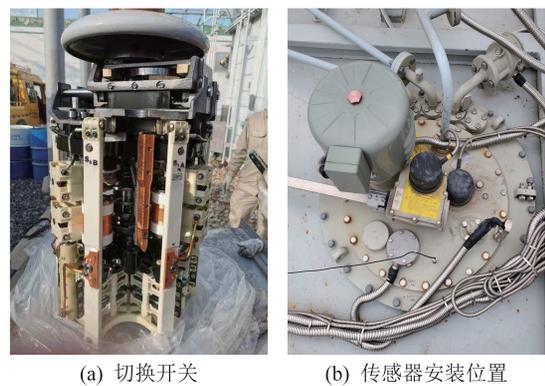


图 7 切换开关和加速度传感器安装位置

Fig.7 Switcher and mounted position of accelerometer

文中对 OLTC 数据集共进行 5 组实验,每组实验中抽取部分(分别为 40 个、80 个、120 个、160 个、200 个)训练样本进行标记,用于监督学习,所有的有标签和无标签样本都会用于无监督学习。每组实验都会对 ConvLN、阶梯网络、BO-ConvLN、CNN、

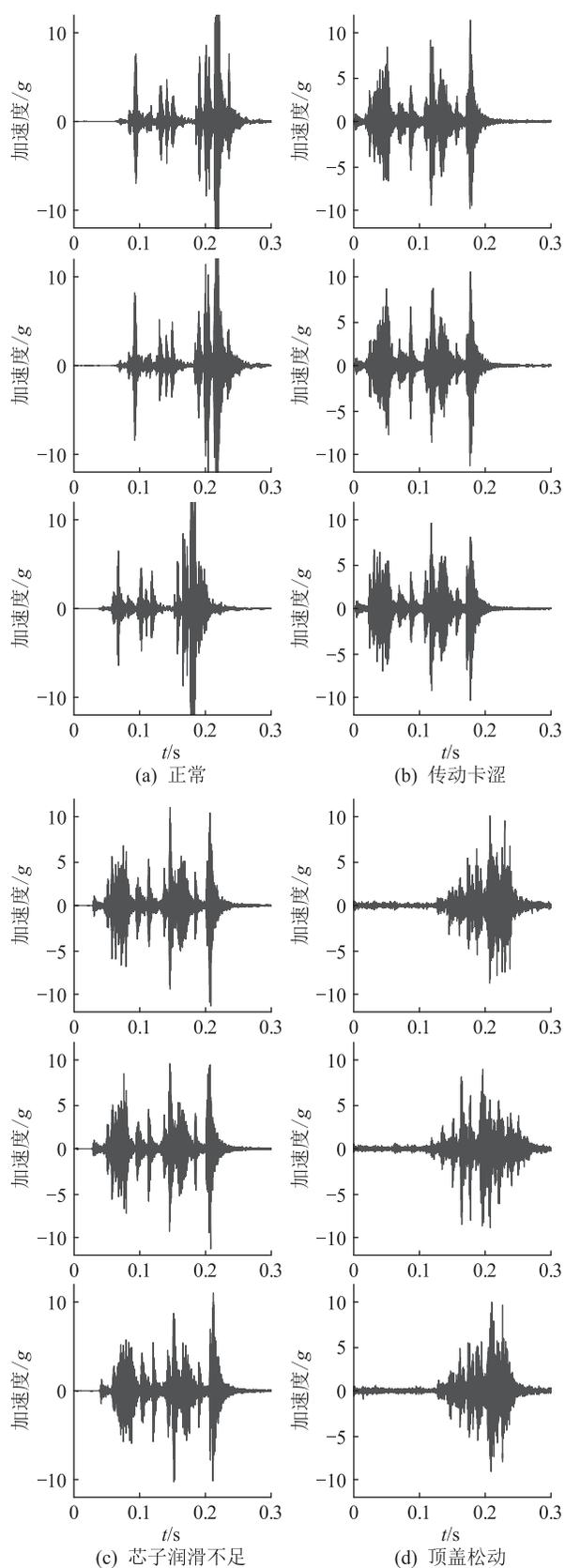


图8 振动信号主波形

Fig.8 Main waveforms of vibration signals

区间最大功率特征(interval maximum power feature, IMPF)+SVM^[10]5类模型进行训练,并借鉴交叉验证的形式每组进行6轮实验,以6轮精度的平均值

作为最终结果。前4类模型为深度学习模型,其参数如下。

ConvLN 编码器和解码器采用对称的卷积和反卷积结构,每层参数分别为:Conv1 为 $10 \times 4 \times 2$ (卷积核数 \times 卷积核长度 \times 步长),Max-pooling2 为 4×4 ,Conv3 为 $10 \times 4 \times 2$,Max-pooling4 为 4×2 ,Conv5 为 $0 \times 4 \times 2$,Max-pooling6 为 4×2 ,Conv7 为 $10 \times 4 \times 2$,Max-pooling8 为 4×4 ,FC9 为 $2\ 240 \times 4$; $\lambda_1 - \lambda_7$ 分别为 10、1、0.1、0.1、0.1、0.1、0.1, A_{noise} 为 0.6;训练时采用 Adam 优化算法,学习率为 0.02(从第 15 代开始衰减),迭代次数为 1 500。阶梯网络编码器和解码器采用对称的 7 层全连接结构,每层神经元数为 30 000、3 000、1 000、1 000、500、100、4,训练方法与 ConvLN 相同。BO-ConvLN 的网络结构与阶梯网络相同, $\lambda_1 - \lambda_7$ 优化范围为 $0.1\lambda \sim 10\lambda$, A_{noise} 优化范围为 0.2~1.5,训练方法与 ConvLN 相同,贝叶斯优化时采用 20 次 Sobol 采样和 20 次 EI 优化提升。CNN 的网络结构与阶梯网络编码器相同,训练时采用 Adam 优化算法,学习率为 0.001,迭代次数为 400。其中 ConvLN 和阶梯网络的超参数 $\lambda_1 - \lambda_7$ 、 A_{noise} 通过手动调参得到。最后一类模型基于人工设计特征 IMPF+传统机器学习模型 SVM 的模式。

在标签数为 200 的情况下,BO-ConvLN 超参数的贝叶斯优化过程如图 9 所示。其中,前 20 次为 Sobol 采样过程,后 20 次为 EI 优化过程。由图 9 可知,EI 优化过程提升了 ConvLN 在验证集上的精度,表明贝叶斯优化方法在较少的迭代次数内能有效优化 ConvLN 的超参数。

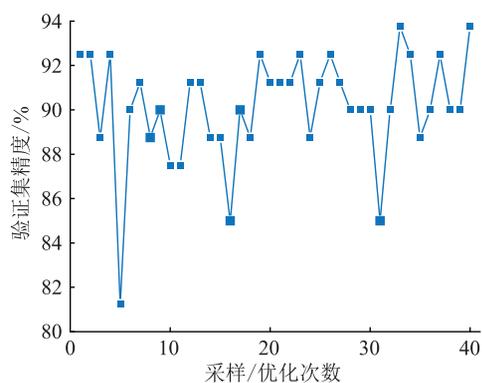


图9 BO-ConvLN 模型的贝叶斯优化过程

Fig.9 Bayesian optimization process of BO-ConvLN model

5类模型在 OLTC 数据集上的实验结果如图 10 所示。

由图 10 可知,在网络结构相同的情况下,由于 ConvLN 融合了无标签样本的额外信息,在 OLTC 数据集上的平均诊断精度比 CNN 高 5.75%。在标签

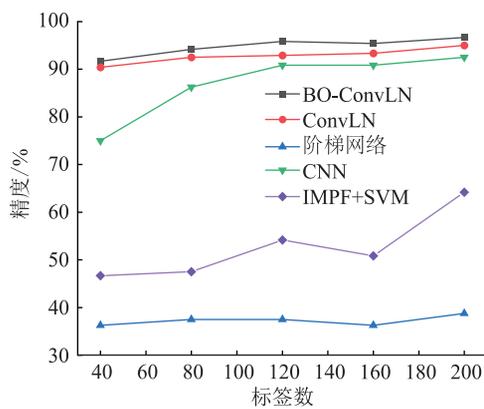


图 10 OLTC 数据集实验结果

Fig.10 Experiment results of OLTC dataset

数为 40 的情况下,ConvLN 的诊断精度远高于 CNN,表明无监督学习在标签数量较少时能显著提升神经网络的诊断精度;在标签数为 200 的情况下(训练数据中所有样本都被标记,ConvLN 与 CNN 训练样本数量相同),ConvLN 的诊断精度仍比 CNN 高 2.5%,表明即使在有标签样本充足的情况下,引入无监督学习仍能提升神经网络的诊断精度。同时,阶梯网络的诊断精度远低于其他 3 类具有卷积算子的模型,这充分说明了对于 OLTC 振动信号这种采样点多、非平稳的信号,卷积算子相较全连接结构具有更强的特征提取能力。此外,通过贝叶斯优化算法对 ConvLN 的超参数进行优化,可提高约 2% 的诊断精度。在标签数仅为 40 的情况下 BO-ConvLN 的诊断精度高达 91.67%,略高于 ConvLN,远高于 CNN 和阶梯网络;在标签数为 200 的情况下,BO-ConvLN 的诊断精度高达 95.51%,略高于 ConvLN 和 CNN,远高于阶梯网络;且 BO-ConvLN 的诊断精度高于传统机器学习模型 IMPF+SVM,证明了在 OLTC 故障诊断领域 BO-ConvLN 相较人工提取特征的模式具有一定优势。

4 结论

文中提出基于贝叶斯优化阶梯网络的 OLTC 故障诊断方法,经过研究和实验,得出以下结论:

(1) 在存在大量无标签样本的情况下,引入半监督学习会显著提高机械故障的诊断精度。在 IMS 数据集上,半监督阶梯网络的平均精度比 CNN 高 14.29%;在 OLTC 数据集上,ConvLN 的平均精度比 CNN 高 5.75%。即使在所有样本都有标签的情况下,ConvLN 仍能带来约 2.5% 的精度提升。

(2) 相较于常见的半监督学习模型 VAE M1、VAE M2、自编码器等,阶梯网络具有更高的诊断精度。基于贝叶斯优化理论对阶梯网络的超参数进

行优化,可进一步提升诊断精度。优化后的超参数在 IMS 数据集上能带来约 3.2% 的精度提升,在 OLTC 数据集上能带来约 2% 的精度提升。

(3) 经过卷积化改造的 BO-ConvLN 对 OLTC 传动卡涩、芯子润滑不足、顶盖松动这 3 类故障的综合诊断精度超过了 90%;在有标签样本数量为 200 时,其诊断精度达到了 95.51%。

致谢

本文得到江苏省科协青年科技人才托举工程项目(2021031)、东南大学“至善青年学者”支持计划项目、中央高校基本科研业务费专项资金项目资助,谨此致谢!

参考文献:

- [1] 朱英浩,沈大中. 有载分接开关电气机理[M]. 北京:中国电力出版社,2012.
ZHU Yinghao, SHEN Dazhong. Electrical mechanism of on-load tap changer[M]. Beijing: China Electric Power Press, 2012.
- [2] 王蕾,袁洪跃,王季琴,等. 变压器有载分接开关技术和故障诊断发展现状及展望[J]. 高压电器,2022,58(3):171-180.
WANG Lei, YUAN Hongyue, WANG Jiqin, et al. Development status and prospect of transformer on-load tap-changer technology and fault diagnosis[J]. High Voltage Apparatus, 2022, 58(3):171-180.
- [3] 张德明. 变压器分接开关保养维修技术问答[M]. 北京:中国电力出版社,2013.
ZHANG Deming. Technical questions and answers of transformer tap changer maintenance[M]. Beijing: China Electric Power Press, 2013.
- [4] 马勇,王树刚,王同磊,等. 基于驱动电机电流与振动信号的有载分接开关故障诊断方法[J]. 高压电器,2022,58(5):202-210.
MA Yong, WANG Shugang, WANG Tonglei, et al. Fault diagnosis method for on-load tap changer based on driven motor current and vibration signal[J]. High Voltage Apparatus, 2022, 58(5):202-210.
- [5] 马勇,王同磊,吴鹏,等. 换流变压器真空有载分接开关动力学模拟及工作一致性研究[J]. 高压电器,2022,58(7):199-206.
MA Yong, WANG Tonglei, WU Peng, et al. Study on the dynamic simulation and working consistency of vacuum on-load tap changer in converter transformers[J]. High Voltage Apparatus, 2022, 58(7):199-206.
- [6] 赵书涛,王紫薇,胡经纬,等. 基于振动信号分析的有载分接开关故障诊断研究综述[J]. 华北电力大学学报(自然科学版),2021,48(5):61-71.
ZHAO Shutao, WANG Ziwei, HU Jingwei, et al. Review on fault diagnosis of on-load tap changer based on vibration signal analysis[J]. Journal of North China Electric Power University (Natural Science Edition), 2021, 48(5):61-71.
- [7] 周志华. 机器学习[M]. 北京:清华大学出版社,2016.

- ZHOU Zhihua. Machine learning [M]. Beijing: Tsinghua University Press, 2016.
- [8] 张知先, 陈伟根, 汤思蕊, 等. 基于互补集总经验模态分解和局部异常因子的有载分接开关状态特征提取及异常状态诊断[J]. 电工技术学报, 2019, 34(21): 4508-4518.
- ZHANG Zhixian, CHEN Weigen, TANG Sirui, et al. State feature extraction and anomaly diagnosis of on-load tap-changer based on complementary ensemble empirical mode decomposition and local outlier factor[J]. Transactions of China Electrotechnical Society, 2019, 34(21): 4508-4518.
- [9] 张伟政, 施琳琛, 汲胜昌, 等. 小波奇异性检测诊断有载分接开关故障[J]. 高电压技术, 2006, 32(7): 49-53.
- ZHANG Weizheng, SHI Linchen, JI Shengchang, et al. Wavelet singularity detection used for fault diagnosis of on-load tap changers[J]. High Voltage Engineering, 2006, 32(7): 49-53.
- [10] 段若晨, 王丰华, 周荔丹, 等. 利用窄带噪声辅助多元经验模态分解算法检测换流变压器用有载分接开关机械状态[J]. 电工技术学报, 2017, 32(10): 182-189.
- DUAN Ruochen, WANG Fenghua, ZHOU Lidan, et al. Mechanical condition detection of on-load tap-changer in converter transformer based on narrowband noise assisted multivariate empirical mode decomposition algorithm[J]. Transactions of China Electrotechnical Society, 2017, 32(10): 182-189.
- [11] 王丰华, 曾全昊, 郑一鸣, 等. 基于 Bayes 估计相空间融合和 CM-SVDD 的有载分接开关机械故障诊断[J]. 中国电机工程学报, 2020, 40(1): 358-368, 402.
- WANG Fenghua, ZENG Quanhao, ZHENG Yiming, et al. A mechanical fault diagnosis of on-load tap-changers based on phase space fusion of Bayes estimation and CM-SVDD [J]. Proceedings of the CSEE, 2020, 40(1): 358-368, 402.
- [12] 赵彤, 李庆民, 陈平. OLTC 振动信号特征提取的动力学分析方法[J]. 电工技术学报, 2007, 22(1): 41-46.
- ZHAO Tong, LI Qingmin, CHEN Ping. Dynamic analysis method for feature extraction of mechanical vibration signals of on-load tap changers[J]. Transactions of China Electrotechnical Society, 2007, 22(1): 41-46.
- [13] 曾全昊, 王丰华, 郑一鸣, 等. 基于卷积神经网络的变压器有载分接开关故障识别[J]. 电力系统自动化, 2020, 44(11): 144-151.
- ZENG Quanhao, WANG Fenghua, ZHENG Yiming, et al. Fault recognition of on-load tap-changer in power transformer based on convolutional neural network [J]. Automation of Electric Power Systems, 2020, 44(11): 144-151.
- [14] LIANG X H, WANG Y Y, GU H R. A mechanical fault diagnosis model of on-load tap changer based on same-source heterogeneous data fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-9.
- [15] 杨祎, 崔其会, 丁奕齐. 面向电网设备故障报告的半监督命名实体识别方法[J]. 计算机应用, 2021, 41(S2): 41-47.
- YANG Yi, CUI Qihui, DING Yiqi. Named entity recognition method for power grid equipment fault report based on semi-supervised learning[J]. Journal of Computer Applications, 2021, 41(S2): 41-47.
- [16] 于希娟, 孙宏伟. 基于图像处理和半监督学习的变电设备故障诊断[J]. 电网与清洁能源, 2022, 38(8): 60-68.
- YU Xijuan, SUN Hongwei. Fault diagnosis of substation equipment based on image processing and semi-supervised learning [J]. Power System and Clean Energy, 2022, 38(8): 60-68.
- [17] PEZESHKI M, FAN L X, BRAKEL P, et al. Deconstructing the ladder network architecture [C]//Proceedings of the 33rd International Conference on Machine Learning-Volume 48. New York, NY, USA. New York: ACM, 2016: 2368-2376.
- [18] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the 32nd International Conference on Machine Learning-Volume 37. New York: ACM, 2015: 448-456.
- [19] RASMUS A, VALPOLA H, HONKALA M, et al. Semi-supervised learning with ladder networks [EB/OL]. [2022-09-24]. <https://arxiv.org/abs/1507.02672>.
- [20] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning [M]. Cambridge, Massachusetts: The MIT Press, 2016.
- [21] 高春永, 柏业超, 王琼. 基于改进的半监督阶梯网络 SAR 图像识别[J]. 南京大学学报(自然科学), 2021, 57(1): 160-166.
- GAO Chunyong, BAI Yechao, WANG Qiong. SAR image recognition based on improved semi-supervised ladder network [J]. Journal of Nanjing University (Natural Science), 2021, 57(1): 160-166.
- [22] 崔佳旭, 杨博. 贝叶斯优化方法和应用综述[J]. 软件学报, 2018, 29(10): 3068-3090.
- CUI Jiayu, YANG Bo. Survey on Bayesian optimization methodology and applications [J]. Journal of Software, 2018, 29(10): 3068-3090.
- [23] MEI J C, ZHANG G J, QI D L, et al. Accelerated solution of the transmission maintenance schedule problem: a Bayesian optimization approach [J]. Global Energy Interconnection, 2021, 4(5): 493-500.
- [24] 何志昆, 刘光斌, 赵曦晶, 等. 高斯过程回归方法综述[J]. 控制与决策, 2013, 28(8): 1121-1129, 1137.
- HE Zhikun, LIU Guangbin, ZHAO Xijing, et al. Overview of Gaussian process regression [J]. Control and Decision, 2013, 28(8): 1121-1129, 1137.
- [25] FRAZIER P I. A tutorial on Bayesian optimization [EB/OL]. [2022-09-24]. <https://arxiv.org/abs/1807.02811>.
- [26] 龚雪娇, 朱瑞金, 唐波. 基于贝叶斯优化 XGBoost 的短期峰值负荷预测[J]. 电力工程技术, 2020, 39(6): 76-81.
- GONG Xuejiao, ZHU Ruijin, TANG Bo. Short-term peak load forecasting based on Bayesian optimization XGBoost [J]. Electric Power Engineering Technology, 2020, 39(6): 76-81.
- [27] 孙琪, 于永进, 王玉彬, 等. 采用改进鲸鱼算法的配电网综合优化[J]. 电力系统及其自动化学报, 2021, 33(5): 22-29.
- SUN Qi, YU Yongjin, WANG Yubin, et al. Comprehensive optimization of distribution network using improved whale optimization algorithm [J]. Proceedings of the CSU-EPSA, 2021, 33

(5):22-29.

- [28] LEE J, QIU H, YU G, et al. Bearing dataset[EB/OL]. [2022-09-24]. <https://ti.arc.nasa.gov/tech/dash/groups/pcoc/prognostic-data-repository/>.
- [29] ZHANG S, YE F, WANG B N, et al. Semi-supervised bearing fault diagnosis and classification using variational autoencoder-based deep generative models [J]. IEEE Sensors Journal, 2021, 21(5):6476-6486.
- [30] BERGSTRA J, BENGIO Y. Random search for hyper-parameter optimization [J]. Journal of Machine Learning Research, 2012, 13:281-305.

作者简介:



郑尚直

郑尚直(1997),男,硕士在读,研究方向为人工智能在电力系统中的应用(E-mail: zhengshangzhi@seu.edu.cn);

仲林林(1990),男,博士,副研究员,博士生导师,研究方向为高电压技术、放电等离子体技术、人工智能技术;

王同磊(1990),男,博士,高级工程师,从事电力设备状态检测、健康评估、故障诊断技术工作。

Fault diagnosis method for OLTC based on improved semi-supervised ladder networks

ZHENG Shangzhi¹, ZHONG Linlin¹, WANG Tonglei², GAO Bingtuan¹

(1. School of Electrical Engineering, Southeast University, Nanjing 210096, China;

2. Electric Power Research Institute of State Grid Jiangsu Electric Power Co., Ltd., Nanjing 211103, China)

Abstract: On-load tap changers (OLTC) have complex mechanical and electrical structures, which are the key component for the on-load voltage regulation of transformers. Currently, due to the sample data which are not easy to be labeled, it is difficult to effectively train the OLTC mechanical fault diagnosis models based on vibration signals. To improve the fault diagnostic accuracy for OLTC with limited labeled data, a fault diagnosis method based on Bayesian optimization-convolutional ladder networks (BO-ConvLN) is proposed in this paper. Firstly, the ladder networks are used as a semi-supervised learning method for the feature extraction of vibration signals, which is guided by a large amount of unlabeled data, leading to the enhanced diagnostic ability of ladder networks only with a small amount of labeled data. Then, the fully-connected layers are replaced by convolutional operators in the ladder networks to better extract the features of non-stationary vibration signals. Furthermore, Bayesian optimization is used to optimize the high-dimensional hyperparameters of ladder networks, which significantly improves the diagnostic accuracy of the model within limited time cost. The experiment results show that the diagnostic accuracy for the three types of faults, namely transmission shaft jams, poor switch lubrication, and top cover looseness, is 91.67% with a label count of only 40, which demonstrates the effectiveness of BO-ConvLN in the fault diagnosis.

Keywords: on-load tap changer (OLTC); fault diagnosis; vibration signals; ladder networks; semi-supervised learning; Bayesian optimization

(编辑 陆海霞)