

DOI: 10.19650/j.cnki.cjsi.J2209225

基于连续小波变换和 MTL-SEResNet 的 断路器故障程度评估*

孙曙光², 张婷婷², 王景芹¹, 魏 硕², 邵 旭²

(1. 河北工业大学省部共建电工装备可靠性与智能化国家重点实验室 天津 300130;

2. 河北工业大学人工智能与数据科学学院 天津 300130)

摘要:考虑到万能式断路器触头系统机械故障是一个从轻微到重度的演变过程,准确识别其运行状态可以大大提高断路器的可靠性。提出一种单信号输入和多任务输出的 MTL-SEResNet 网络模型以兼顾故障诊断和程度评估。首先采用连续小波变换对触头系统振动信号进行时频分析,得到相应的二维时频图像;其次将 SENet 结构引入到改进的 ResNet18 中,利用多任务学习共享机制构建 MTL-SEResNet 网络模型;并通过调整故障分类和程度评估两个任务损失函数的权重比例,对模型进行优化;最后,通过模拟的触头系统的故障数据对所提方法进行实验验证。结果表明,模型的性能更佳,类型及程度准确率分别为 99.78% 和 99.36%,可以有效地实现万能式断路器故障程度评估。

关键词: 万能式断路器;触头系统;故障程度评估;连续小波变换;多任务学习

中图分类号: TM561 TH165.3 **文献标识码:** A **国家标准学科分类代码:** 470.40

Fault degree evaluation of circuit breaker based on continuous wavelet transform and MTL-SEResNet

Sun Shuguang², Zhang Tingting², Wang Jingqin¹, Wei Shuo², Shao Xu²

(1. State Key Lab Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300130, China;

2. School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China)

Abstract: The mechanical fault of the contact system for a conventional circuit breaker is a process from slight to severe. The accurate identification of its operating state can greatly improve the reliability of the circuit breaker. In this article, a single signal input and multi-task output MTL-SEResNet model is proposed for fault diagnosis and degree evaluation. Firstly, the raw vibration signals of the contact system are analyzed using a continuous wavelet transform. And the corresponding two-dimensional time-frequency images are obtained. Secondly, the improved ResNet18 network is combined with the SENet structure, and the multi-task learning sharing mechanism is used to formulate the MTL-SEResNet model. The model is optimized by adjusting the weight ratio of the two task loss functions for fault classification and degree evaluation. Finally, the proposed method is verified by experiments with simulated fault data of the contact system. The results show that the proposed model has better performance with 99.78% and 99.36% accuracy in type and degree, respectively, which can effectively evaluate the fault degree of the conventional circuit breaker.

Keywords: conventional circuit breaker; contact system; fault degree evaluation; continuous wavelet transform; multi-task learning

0 引 言

万能式断路器主要用于低压配电网中,其处于运行状态时可能发生各种故障,不仅会产生巨大的财产损失,而且还会对操作人员以及整个配电网的安全造成巨大威

胁^[1]。断路器在分合闸过程中,动、静触头发生激烈碰撞,产生的振动信号蕴含着触头系统及其操作机构丰富的信息,能够表征断路器的多种运行状态,对其机械状态的变化具有较高的敏感性,并且具有非侵入式的特点^[2-4]。考虑到断路器的机械故障是一个从轻微到重度的演变过程,因此,检测振动信号对万能式断路器进行故

收稿日期:2022-01-19 Received Date: 2022-01-19

* 基金项目:河北省自然科学基金(E2021202136)、河北省自然科学基金创新群体(E2020202142)项目资助

障诊断和程度评估,对确保整个低压配电网的安全可靠运行具有重要现实意义。

利用振动信号进行机械故障诊断,相关学者已经进行了较为深入的研究。其中,传统机器学习方法一般通过变分模态分解、小波包分解和局部均值分解等方法对原始振动信号进行模态分解,再提取特征,进而通过分类算法实现故障诊断^[5-6]。Shao等^[7]通过将变分模态分解、时移多尺度离散熵和支持向量机(support vector machine, SVM)结合在一起,实现了轴承故障诊断。文献^[8]提出了一种基于加权极限学习机的小波包分解和核主成分分析的故障诊断方法。文献^[9]利用局部均值分解和多尺度排列熵等方法,实现了万能式断路器的故障诊断和程度定量评估。然而这些方法比较依赖于专家知识,当信号结构复杂时,难以提取到有效故障特征。

与传统机器学习方法不同,深度学习模型具有更强的非线性特征学习和故障特征分析能力^[10-13]。利用深度学习模型进行故障诊断时,对于振动信号的处理一般采取两种方式:1)直接输入原始振动信号;2)对振动数据通过时频分析方法处理后输入模型。直接输入原始振动信号,由于分析域的局限,并不能全面反映机械故障的状态信息^[14]。振动信号具有非线性和非平稳性的特点,分析这种信号常用的时频分析法有短时傅里叶变换(short-time Fourier transform, STFT)^[15]、连续小波变换(continuous wavelet transform, CWT)^[16]和希尔伯特-黄变换(Hilbert-Huang transform, HHT)^[17]等。康守强等^[18]对振动信号进行STFT,并引入Resnet18网络对特征进行深层提取。但非平稳信号时间和频率之间存在不可调和的矛盾,导致STFT难以兼顾时域和频域的分辨率;而CWT具有多分辨率特性,能够很好地在时频域内刻画出信号的局部信息。Guo等^[19]对采集到的振动信号进行CWT,以时频图为输入构建卷积神经网络(convolutional neural network, CNN)故障诊断模型。此外HHT用来分析非平稳信号具有自适应性的特点。Yang等^[20]采用HHT对振动信号进行信号处理,并通过三维CNN实现故障诊断。因此针对万能式断路器,应该结合其振动信号的特点,研究更合适的信号处理方法。此外上述研究中,虽然CNN在以时频图为输入的故障诊断任务中已经取得了不错的效果,但是随着CNN层数的增加,可能会发生过拟合现象,导致模型性能发生下降。为了解决这个问题,Zhao等^[21]提出了一种基于深度残差网络(deep residual network, ResNet)故障诊断方法,提高了特征识别能力。田科位等^[22]将SENet网络引入到ResNet的残差块之中,达到了提升模型性能的目的。

由于故障严重程度是在机械发生故障后缓慢积累而成的,故一种有效的故障诊断方法应该包含故障分类和程度评估两个方面。在故障类型和故障严重程度识别方面

已经进行了一些研究。Guo等^[23]利用基于分层学习速率的自适应深度CNN,实现了轴承故障诊断和严重程度判定。Gan等^[24]设计了一种基于小波包能量特征和深度置信网络(deep belief network, DBN)的两级诊断模型。上述方法是一种分层模型,训练了一个用于故障分类的网络和几个用于单一故障状态下程度评估的网络,先识别故障类型再进行故障程度评估。事实上故障类型和程度之间是相互关联的,可以通过一些共享特征进行关联。由于多任务学习(multi-task learning, MTL)可以利用不同任务之间相互关联的信息来改进模型的泛化能力^[25]。Wang等^[26]提出了一种新的多任务注意CNN,用于轴承的故障诊断和工况识别。Wen等^[27]利用STFT和LeNet-5网络,构建了一种用于机械故障分类和程度评估的算法。上述方法通过任务之间的信息迁移,使模型学习效率得到有效提升。

基于上述研究,本文提出了一种基于CWT和MTL-SEResNet网络的万能式断路器触头系统故障程度评估方法。对于原始振动信号利用CWT获取二维时频图作为评估模型的输入;针对评估模型,通过引入SENet结构以及将平均池化(average-pooling, Avgpool)改为全局平均池化层(global average pooling, GAP),对其核心网络模块ResNet18进行改进,在此基础上利用MTL的共享机制实现模型对故障分类和程度评估任务的互助学习,并对模型的损失函数权重比例进行优化,以进一步提升断路器故障分类和程度评估的准确率。

1 万能式断路器故障程度评估方法

本文提出一种涉及振动信号的CWT以及基于MTL共享机制的MTL-SEResNet网络模型的方法。

1.1 连续小波变换

CWT是继STFT之后被认为是处理非平稳信号的又一重大发现,其可以根据不同信号的特点调节窗口大小,进行多分辨率分析。CWT定义如下:

$$W_f(s, \tau) = \langle f, \varphi_{s, \tau}(t) \rangle = |s|^{-1/2} \int f(t) \overline{\varphi\left(\frac{t-\tau}{s}\right)} dt \quad (1)$$

式中: $\overline{\varphi(t)}$ 表示 $\varphi(t)$ 的复共轭;符号 $\langle f, \varphi_{s, \tau} \rangle$ 表示两者的内积, $\varphi_{s, \tau}(t)$ 为小波基函数, s, τ 分别为尺度因子和平移因子。

1.2 MTL-SEResNet 网络模型

1) 多任务学习

MTL是在传统单任务学习基础之上提出的一种机器学习概念,把多个相关的任务放在一起学习,并且其包括Soft和Hard两种共享机制。其中Soft共享机制是指每个任务都有单独的参数和模型,且模型参数之间的距离是正则化的,以便鼓励参数相似化。而Hard共享机制则是指多

个任务通过共享层来互相帮助学习,可提升模型整体的泛化效果,适合处理相关性较强的任务,而本文对万能式断路器触头系统机械故障及其严重程度进行评估,二者本身就存在着较强的内在联系,本文模型采用 Hard 共享机制。

2) ResNet 网络

2016 年,He 等^[28]提出了建立在 CNN 基础上的 ResNet 网络,通过增加残差块解决了深层网络结构的梯度消失问题。在 ResNet 网络结构中,卷积层(convolutional layer, Conv)由多组平行的卷积核组成,卷积核用来计算不同的特征图。并选用 ReLU 激活函数引入非线性特性。故 Conv 的计算公式为:

$$x_j^m = f\left(\sum_{i=0}^n x_i^{m-1} k_{ij}^m + b_j^m\right) \quad (2)$$

$$f\left(\sum_{i=0}^n x_i^{m-1} k_{ij}^m + b_j^m\right) = \max\left\{\sum_{i=0}^n x_i^{m-1} k_{ij}^m + b_j^m, 0\right\} \quad (3)$$

式中: x_i^{m-1} 和 x_j^m 分别为第 m 个卷积层的输入和输出; j 为维度; n 为输入特征向量的数量; k_{ij}^m 和 b_j^m 分别为卷积核和偏置; $f(\cdot)$ 是激活函数。

池化层是对卷积层中提取的特征进行挑选,可以缩减模型的大小,提高计算速度。常见的池化操作有平均池化和最大值池化(max-pooling, Maxpool),计算过程分别如下:

$$p_j^m = \max_{(j-1)l+1 \leq t \leq jl} \{x_j^m(t)\} \quad (4)$$

$$p_j^m = \frac{1}{l} \sum_{(j-1)l+1}^{jl} x_j^m \quad (5)$$

式中: $x_j^m(t)$ 为第 t 个神经元输出的激活值; p_j^m 为池化后的输出; l 为池化层的尺寸。

批量归一化(batch normalization, BN)层是用来解决深层网络训练可能带来的“梯度弥散或梯度爆炸”问题而提出的一种归一化方法。BN 层的计算过程为:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (6)$$

$$v^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (7)$$

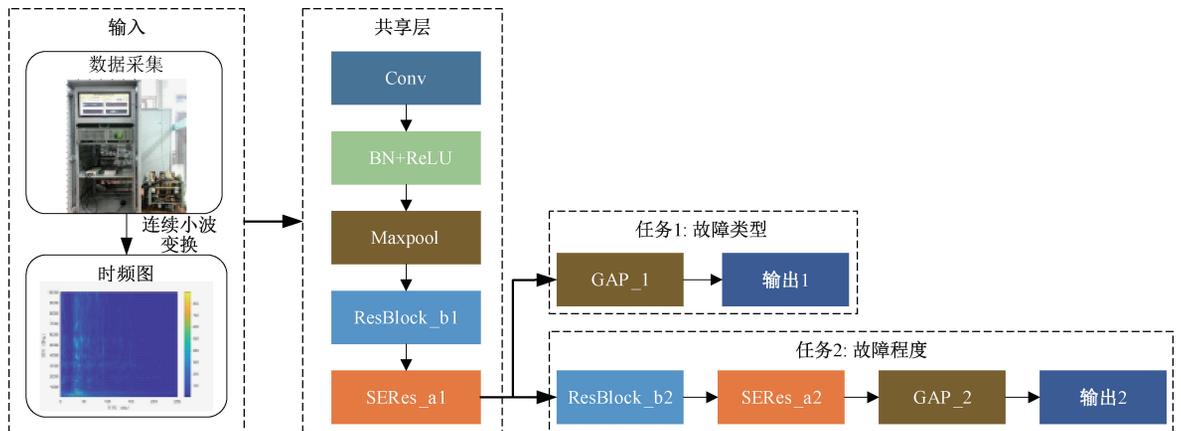


图1 MTL-SEResNet 模型结构

Fig. 1 Network of the MTL-SEResNet model

$$z_i = \frac{x_i - \mu}{\sqrt{v^2 + \epsilon}} \quad (8)$$

$$y_i = \gamma z_i + \beta \quad (9)$$

式中: x_i 和 y_i 分别表示批处理中第 i 个观测值的输入和输出特征; N 是任务中每个批次样本的个数; γ 和 β 是两个可训练的参数; ϵ 是接近 0 的常量。

3) 模型网络结构设计

根据模型的输入和数据集大小,对 ResNet18 进行如下改进:第 1 层卷积核改为 3×3 , Maxpool 尺寸改为 2×2 ,并将 Avgpool 改为 GAP 层;将 SENet 加入到残差块中,组成新的 SERes_a 结构以替换该残差块。其中 SENet 结构^[29],包括挤压(squeeze)和激励(excitation)两大操作,挤压操作对输入的特征图进行全局池化,计算过程如下:

$$F_{sq}(x_q) = \frac{1}{H \times W} \sum_h \sum_w x_q(h, w) \quad (10)$$

式中: x_q 代表输入尺寸为 $H \times W$ 的第 q 个特征图; $F_{sq}(x_q)$ 表示挤压操作过后的输出值。

激励操作由两个全连接层和激活函数组成,计算过程如下:

$$y_q = F_{ex}(F_{sq}(x_q), \omega) = \sigma(\omega_2 \delta(\omega_1 F_{sq}(x_q))) \quad (11)$$

式中: ω_1 为第 1 个全连接层计算; ω_2 为第 2 个全连接层计算; δ 为 ReLU 激活函数; σ 为 Sigmoid 激活函数。

Sigmoid 激活函数的计算公式为:

$$\sigma = \frac{1}{1 + e^{-z}} \quad (12)$$

式中: z 为经过两次全连接计算的输出值。

利用 MTL 的 Hard 共享机制,本文构建的 MTL-SEResNet 网络模型由一个共享层和两个特定任务层组成,如图 1 所示。其中程度评估取 ResNet18 的第 2 和 3 组残差块,故障分类取 ResNet18 的第 2 组残差块,模型参数如表 1 所示。MTL-SEResNet 网络模型中的两种残差块结构如图 2 所示。

表 1 MTL-SEResNet 模型参数

Table 1 MTL-SEResNet model parameters

命名	层	卷积核	步长	卷积核数目
Conv1	卷积	3×3	1×1	32
Max-pooling_1	最大值池化	2×2	2×2	-
ResBlock_b1	ResBlock_b	-	-	64
SERes_a1	SERes_a	-	-	64
ResBlock_b2	ResBlock_b	-	-	128
SERes_a2	SERes_a	-	-	128
GAP1	全局平均池化	2×2	1×1	64
GAP2	全局平均池化	2×2	1×1	128
输出 1	Softmax	-	-	6
输出 2	Softmax	-	-	16

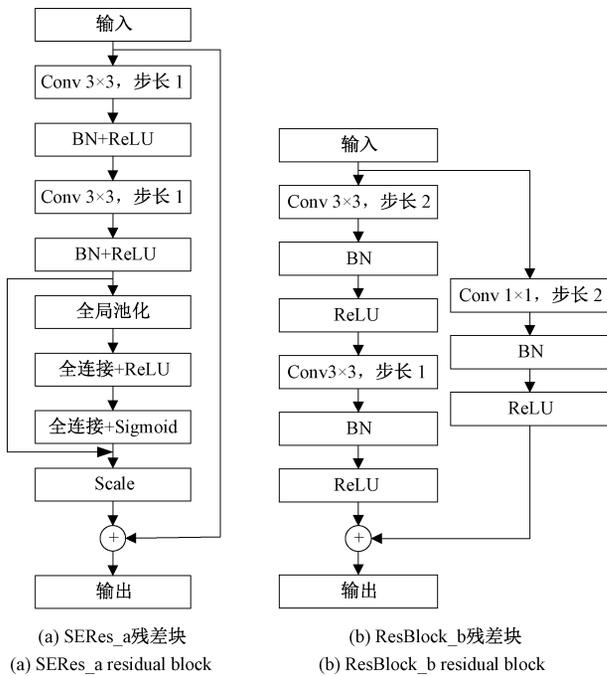


图 2 残差块结构
Fig. 2 Network of residual block

在上述模型中,GAP 是将特征提取层的每个输出特征映射转换为一个值,可以减少过拟合,并降低模型的训练时间。其定义如下:

$$p_j^m = \frac{1}{l} \sum_{l=1}^l x_j^m(t) \quad (13)$$

4) 模型训练与优化

本文模型建立在 TensorFlow2.0 GPU 框架下,电脑配置为 Intel(R) Core(TM) i5-9400F CPU @ 2.90 GHz + NVIDIA GeForce RTX 3060。Batch-size 设置为 32,最大迭代次数为 200 次。模型训练时,采用交叉熵损失函数

计算两个任务的损失加权和,将其作为模型的总损失,计算公式如下:

$$Loss_{总} = aLoss_{类型} + (1 - a)Loss_{程度}, 0 < a < 1 \quad (14)$$

式中: a 为故障类型识别任务损失所占的比重。

假设振动信号经过 Softmax 分类器输出为 $Q(x)$,其目标值为 $P(x)$ 。则式(14)中损失函数为:

$$Loss = - \sum_x (P(x) \lg Q(x) + (1 - P(x)) \lg(1 - Q(x))) \quad (15)$$

由于 Adam 优化器收敛速度快,对内存需求较少,可以降低对设备的性能需求。故采用 Adam 优化器,学习率设置为 0.001,对模型进行优化。并通过对每项任务损失的比重进行调节,使两项任务的训练效果达到最优。

1.3 所提模型的故障程度评估流程

本文所提模型的故障程度评估流程具体步骤如下。

- 1) 获取万能式断路器触头系统在分合闸过程中不同工作状态下的振动信号。
- 2) 对采集到的振动信号进行 CWT,得到转化后的二维时频图。
- 3) 按照一定的比例将经过 CWT 后的数据集随机划分为训练数据和测试数据。
- 4) 建立 MTL-SEResNet 模型,并初始化参数。
- 5) 输入训练数据,并计算模型交叉熵损失函数。
- 6) 判断训练是否达到最大批次,若达到则将测试数据输入到训练好的模型。否则转到步骤 5)。
- 7) 输出 MTL-SEResNet 模型故障程度评估的结果。

2 断路器故障诊断与程度评估实验研究

以模拟的触头系统故障数据为数据集,设置不同的实验模式对所提方法进行测试,并与其他方法进行对比,证明了本文模型的优越性。

2.1 实验装置与数据集

本文研究 DW15-1600 万能式断路器触头系统在分合闸过程中产生的机械故障,包括正常状态,A 相不同期 (NA)、B 相不同期 (NB)、C 相不同期 (NC)、虚假合闸 (FC) 和分闸不彻底 (IO)。实验现场如图 3 所示。图 3 中,LC0159 加速度传感器进行振动信号测量,安装于断路器的横梁,加速度传感器由 LC0201 信号调理器供电,同时隔除偏置电压,最后,通过 USB-7648A 数据采集卡进行振动信号采样。由于断路器机械结构的原因,在进行各故障的程度模拟时,触头的开距与超程只能以 0.5 mm 为单位进行调节,且最大为 5.0 mm。故本文模拟了 0.5、1.0、1.5、2.0、2.5、3.0、3.5、4.0、4.5 mm 共 9 种距离下的 2 160 组故障数据,并且将 2.0 mm 以下设置为轻微故障,2.0~3.0 mm 设置为一般故障,3.5~4.5 mm

设置为重度故障。5 种故障的模拟方式如表 2 所示,相应的调节机构如图 3 所示。其中数据采集卡的采样频率为 20 kHz, 单次振动采样时长为 250 ms。每组样本数据的长度为 5 000 个采样点, 经过 CWT 转换为 128×128 大小的二维图像, 并随机选取 80% 的数据为测试集, 每种工作状态下的样本数据如表 3 所示。



图 3 实验现场

Fig. 3 Experimental site

表 2 故障模拟方式

Table 2 Fault simulation mode

故障类型	模拟方式
A 相不同期	通过连杆调节某相触头与另外两相触头超程大小
B 相不同期	
C 相不同期	
虚假合闸	通过悬臂调节三相触头超程大小
分闸不彻底	在分闸挡板加垫片调整三相触头开距的大小

2.2 断路器振动信号时频分析

正常工作状态以及一般程度下各故障状态的振动信号的时域波形如图 4 所示, 可以看出, 断路器振动信号具有非线性和非平稳性的特点。且不同工作状态之间的时域波形区别不大, 难以直接通过单一的时域分析进行区分。因此需要进一步研究合适的时频分析方法, 在此分别对海明窗窗长 128 的 STFT、基于 cmor3-3 小波的 CWT 和 HHT 3 种方法进行分析, 以正常状态下的振动信号为例构建时频图, 如图 5 所示。时频图中振动信号的频

表 3 实验数据集

Table 3 Experimental data

工作状态	故障程度	类型标签	程度标签	样本数量
正常	-	0	0	360
A 相不同期	轻微	1	1	120
	一般		2	120
	重度		3	120
B 相不同期	轻微	2	4	120
	一般		5	120
	重度		6	120
C 相不同期	轻微	3	7	120
	一般		8	120
	重度		9	120
虚假合闸	轻微	4	10	120
	一般		11	120
	重度		12	120
分闸不彻底	轻微	5	13	120
	一般		14	120
	重度		15	120

率随时间进行变化, 通过冷暖颜色反映各频率成分的能量大小。

由图 5 可知, STFT 时频图能量大的部分呈不规则形状分布, 且沿时间轴方向能量变化十分复杂, 难以同时兼顾时域和频域的分辨率; HHT 时频图的频率波动比较大, 且在 25 ms 左右出现了频率交叉现象; 而 CWT 时频图能量较为集中, 能够刻画出信号时间和频率的局部信息, 相对于上述两种方法具有更好的时频分辨率。

为了进一步证明本文选取 CWT 方法对振动信号进行时频处理的可行性, 分别以 STFT、HHT 和 CWT 处理后 128×128 大小的时频图作为 MTL-SEResNet 模型输入, 进行对比实验。其中模型参数设置如表 1 所示。为了避免随机性的影响, 每种时频图进行 20 次实验, 实验结果如表 4 所示, 其中“99.57 (95.37)”表示故障分类平均准确率为 99.57%, 程度评估平均准确率为 99.37%。

表 4 不同时频图输入

Table 4 Different time-frequency image inputs %

模型输入	平均准确率	
	训练	测试
STFT 时频图	99.57 (99.29)	95.37 (95.63)
HHT 时频图	98.90 (99.93)	96.31 (96.06)
CWT 时频图	100.00 (99.83)	99.78 (99.36)

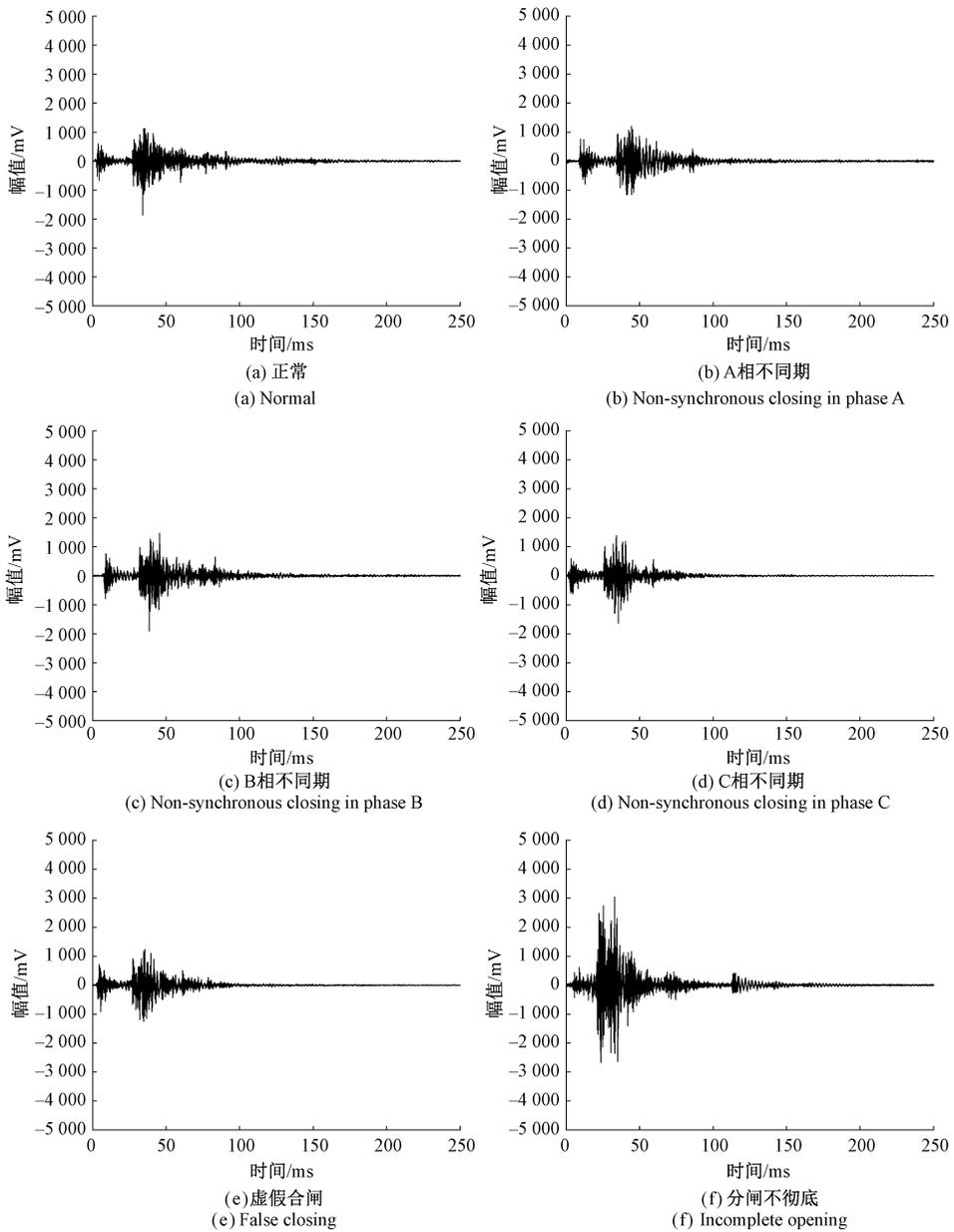


图 4 时域波形

Fig. 4 Time-domain waveform

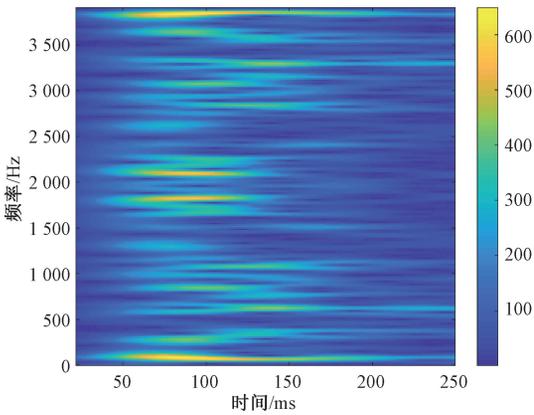
由实验结果可知, 由于 STFT 时频图能量分布较乱, 且 HHT 时频图中波动较大, 使得两种时频图中蕴含的故障特征信息不太明显, 模型存在较明显的过拟合现象。而 CWT 时频图作为输入时, 由于能量分布较为集中, 使得故障分类和程度评估两个任务的平均准确率较其他两种时频图输入提高了 3% 左右。综上所述, 本文选用 CWT 对万能式断路器振动信号进行时频转换, 通过提取 CWT 时频图的特征实现对断路器触头系统故障程度的评估。

2.3 MTL-SEResNet 诊断与评估模型测试

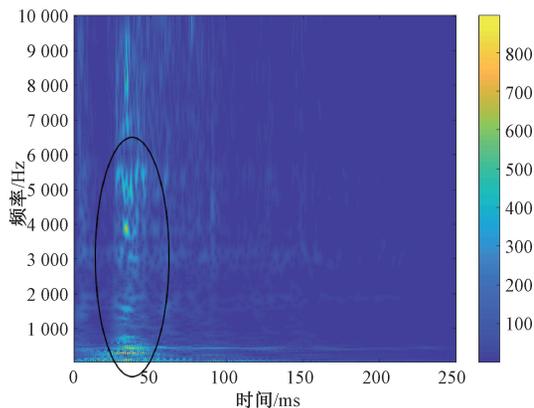
为了评估模型的性能, 本文分别从损失权重的比例、时频图像的大小、网络结构以及卷积核数目 4 个方面对 MTL-SEResNet 模型进行测试。并且以准确率 (accuracy)、召回率 (recall)、精确率 (precision) 为评价指标:

$$accuracy = \frac{1}{C} \sum_{c=1}^C \frac{TP_c + TN_c}{TP_c + TN_c + FP_c + FN_c} \quad (16)$$

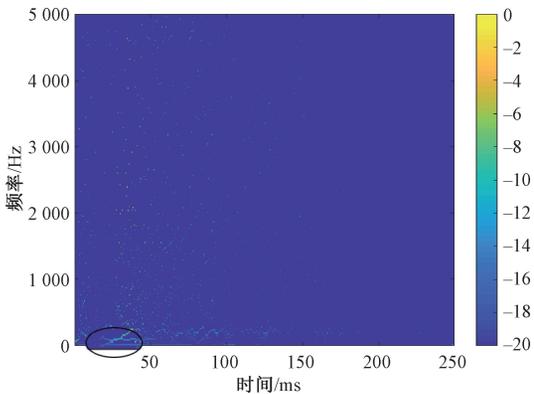
$$recall = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FN_c} \quad (17)$$



(a) 短时傅里叶变换
(a) Short-time Fourier transform



(b) 连续小波变换
(b) Continuous wavelet transform



(c) 希尔伯特-黄变换
(c) Hilbert-Huang transform

图5 振动信号时频图

Fig. 5 Time-frequency image of vibration signal

$$precision = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c} \quad (18)$$

式中: TP 表示正确分类的阳性样本; FP 表示错误分类的阴性样本; FN 表示错误分类的阳性样本; TN 表示准确分类的阴性样本; C 表示模型最终输出的类别。

1) 不同损失权重比例

由于故障诊断和程度评估两个任务的损失比例会影

响模型的诊断结果,故本文选取不同的损失权重比例分别进行实验 20 次,实验结果如图 6 所示。

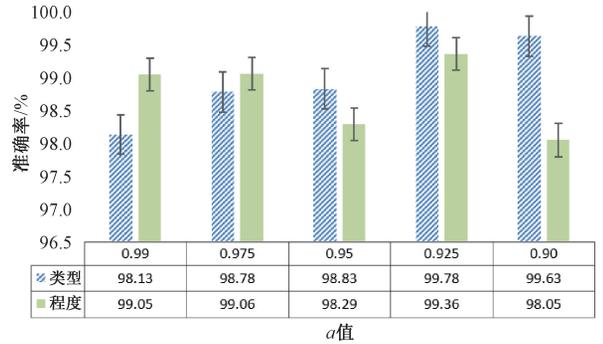


图6 不同损失权重比例

Fig. 6 Different weight ratio

实验结果表明,当 a 取 0.925 时,MTL-SEResNet 模型能够达到较好的评估性能,并且为故障诊断和程度评估的进一步研究提供了基础。

2) 时频图大小

本文采用 cmor3-3 小波将振动信号转换为二维时频图,转换后的图像大小为 605×530。但由于图像大小会对模型的运行时间和评估结果有一定的影响,所以需要时频图进行网格规范化压缩处理。分别选用尺寸大小为 32×32、64×64、128×128、224×224 的图像进行 20 次实验,训练集和测试集的准确率曲线如图 7 所示。图 7 中“train1、test1”表示故障分类任务的训练集、测试集准确率曲线,“train2、test2”表示程度评估任务的训练集、测试集准确率曲线。

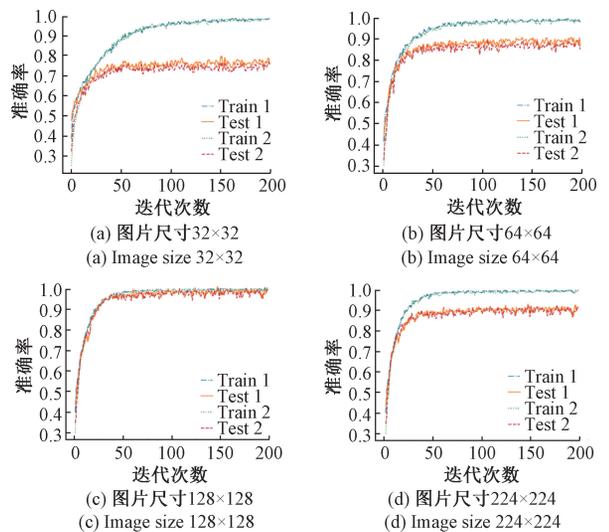


图7 不同图片大小

Fig. 7 Different image sizes

由图 7 可以看出,随着图像尺寸变大,两个任务的测试集准确率先升高然后降低。图片尺寸为 128×128 时,两个任务中的 4 条准确率曲线之间最接近,且趋于平稳后的准确率最高,故本文选取图像尺寸为 128×128。

3) 网络结构

为证明本文所提模型的优势,分别与 ResNet18、ResNet18_1 和 ResNet18_2 进行对比,3 种模型都是采用多任务模型结构。为了测试每种模型的稳定性,分别进行 20 次实验,结果如表 5 所示。由于每个网络有两个分类器,表 5 中“96.50 (93.56)”表示故障分类的结果为 96.50%,程度评估的结果为 93.56%。

表 5 不同网络结构
Table 5 Different network structures

模型	ResNet18	ResNet18_1	ResNet18_2	MTL_SEResNet
准确率/%	96.50 (93.56)	97.75 (97.75)	97.98 (97.84)	99.78 (99.36)
召回率/%	97.49 (93.55)	98.74 (98.74)	98.99 (98.85)	100 (100)
精确率/%	95.51 (92.57)	96.76 (96.76)	96.97 (96.83)	98.57 (98.35)
参数量	11 248 278	718 550	663 638	666 402
运行时间/min	16.67	6.67	6.67	6.67

(1) ResNet18, 程度评估采用原始 ResNet18 结构,而故障分类则去掉 ResNet18 网络中的最后一组残差块。

(2) ResNet18_1, 去除本文所提模型的 SENet 部分,将 GAP 改为 Avgpool,其他部分与本文模型一致。

(3) ResNet18_2, 去除本文所提模型的 SENet 部分,其他部分与本文模型一致。

由对比结果可知,原始 ResNet18 模型参数量最多,使得运行时间较长,且评估效果并不好;将 Avgpool 改为 GAP 后模型的参数量减少,准确率更好;SENet 结构引入 ResNet18 残差块后,模型提取故障特征的能力有所提升。

4) 卷积核数目

由于卷积核的大小会影响模型最后的评估效果,故本文根据消融实验的原理以准确率作为评价指标,分别选取不同的卷积核组合进行 20 次实验。实验结果如表 6 所示,其中“94.28 (93.65)”表示故障分类准确率为 94.28%,程度评估准确率为 93.65%。

由表 6 可以看出,当卷积核数目相对较小时,故障分类和程度评估的准确率都较低,模型整体诊断结果不理想;当卷积核数目过大时,两个任务同时出现了严重的过拟合现象,且模型测试准确率急剧下降;当卷积核组合为 32-64-64-128-128 时,故障分类和程度评估在训练和测试后的准确率相差较小且准确率最高。因此 MTL-SEResNet 模型最终选择卷积核组合为 32-64-64-128-128 对故障分类和程度评估进行研究。

表 6 不同卷积核组合

Table 6 Different convolution kernel combinations

卷积核组合	准确率/%	
	训练	测试
16-32-32-64-64	94.28 (93.65)	92.29 (91.44)
32-32-32-64-64	94.27 (93.36)	92.28 (91.43)
32-64-64-128-128	100.00 (99.83)	99.78 (99.36)
64-64-64-128-128	99.93 (99.93)	93.29 (93.36)
64-128-128-256-256	99.99 (99.89)	94.29 (92.67)

5) 测试结果和可视化分析

在上述研究的基础之上,为更加直观分析本文模型的评估效果,利用已经训练好的模型对测试数据进行识别,最终以混淆矩阵的形式进行输出,如图 8 所示。

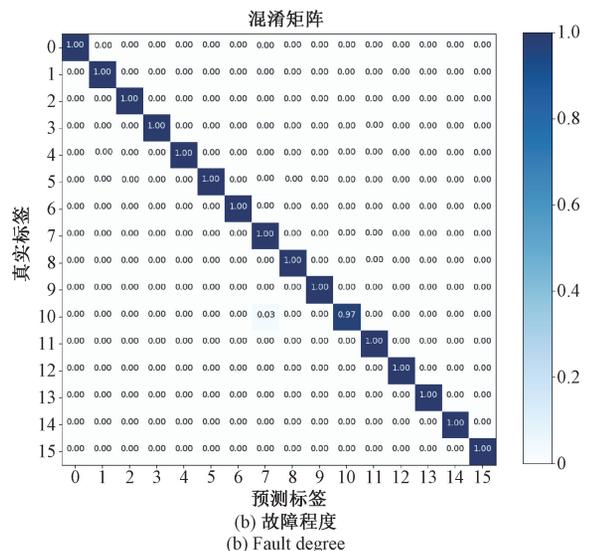
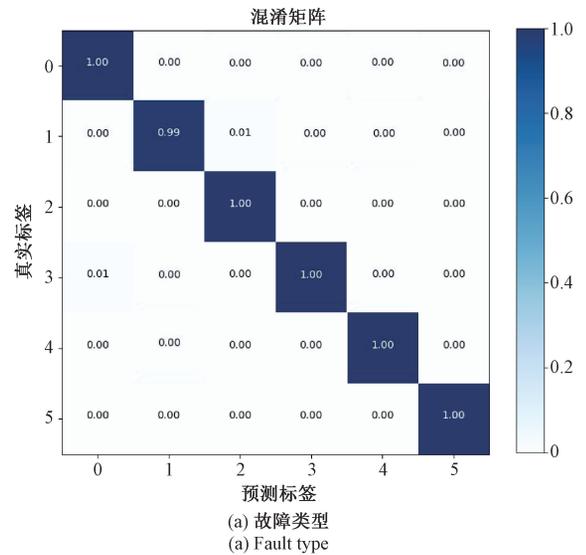


图 8 实验结果

Fig. 8 Experimental results

由故障类型混淆矩阵可以看出,只有 1%的 A 相不同期被错误分类为 B 相不同期。同理从故障程度混淆矩阵中可以看出,只有 3%的虚假合闸轻微故障被错误分类为 C 相不同期轻微故障,表明本文模型不仅可以有效判别故障的类型,同时也可以较准确判断出故障的严重程度。为了更好地理解模型故障分类和程度评估的过程,以 A 相不同期轻微故障时频图为输入,进行特征可视化,如图 9 所示。

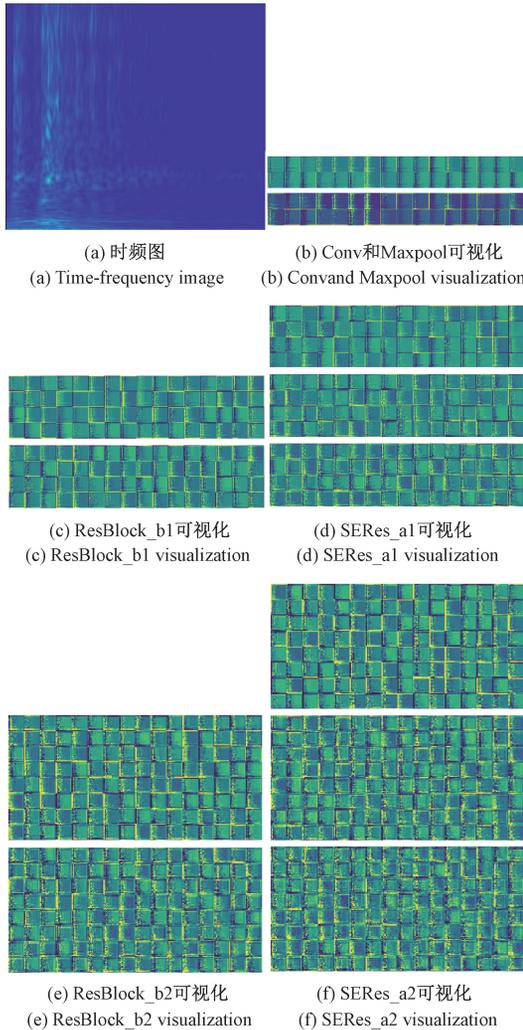


图 9 MTL-SEResNet 模型特征可视化

Fig. 9 MTL-SEResNet model feature visualization

比较 MTL-SEResNet 模型不同层次的特征可以看出,原始图像经过浅层操作之后,提取的是图像的边缘低级特征,由于不同的卷积核对不同的图像内容敏感,随着网络层数的加深,特征图分辨率越来越小,从而提取到越来越抽象的高级特征,各网络分支最终可实现相应的类型诊断以及程度评估任务。

2.4 对比实验

1) 与分层结构模型对比

将 MTL-SEResNet 网络与采取分层结构的 CNN^[23]、DBN^[24]、SVM^[30] 3 种网络进行对比。其中,CNN 直接输入原始振动信号,DBN 输入经小波包变换处理后的振动信号,SVM 以振动信号的时间分割能量熵作为输入,对比结果如表 7 所示。

表 7 分层模型对比结果

Table 7 Comparison results of hierarchical model %

模型	类型	程度					平均准确率
		NA	NB	NC	FC	IO	
本文	99.78	100	100	100	97.64	100	99.57
CNN	96.38	97.05	96.19	96.75	96.01	96.54	96.49
DBN	94.04	94.42	95.01	94.89	93.67	94.53	94.43
SVM	90.84	91.23	90.75	90.14	91.83	92.69	91.25

由表 7 可知,3 种网络结构中,基于浅层网络的 SVM 在准确率上较采用深度学习方法的 CNN、DBN 存在差距,但其中识别率最高的 CNN 由于采用分层结构,且直接通过原始振动信号进行故障程度评估,并不能全面反映断路器故障状态信息,使得准确率低于本文模型;本文所提的 MTL-SEResNet 通过学习二维时频图像的特征,以及借助于网络结构特有的故障类型和程度之间的信息迁移能力,使得平均识别率较其他 3 种网络得到了有效提升。

2) 与多任务学习模型对比

为了验证多任务 Hard 机制下本文模型的优势,设置 HCNN^[27]、ARAE^[31]、FD、DI、MTL-SEResNet 五种网络结构,其中 FD 为 MTL-SEResNet 的故障诊断网络分支结构,DI 为 MTL-SEResNet 的程度评估分支结构。模型输入的原始振动信号长度为 5 000,二维时频图大小为 128×128。对比结果如表 8 所示。

表 8 多任务学习模型对比结果

Table 8 Comparison results of multi-task learning model

模型	输入	准确率/%		运行时间 /min
		类型	程度	
ARAE	原始振动信号	94.89	93.39	10.00
MTL-SEResNet	原始振动信号	97.64	97.32	5.00
HCNN	二维时频图	92.64	81.07	8.33
FD	二维时频图	98.44	-	6.67
DI	二维时频图	-	98.21	6.67
MTL-SEResNet	二维时频图	99.78	99.36	6.67

从实验结果来看,首先,MTL-SEResNet 的各网络分支在面对单一任务时能够取得 98.44% 和 98.21% 准确率,相对于多任务机制下的结果而言也进一步体现了 Hard 多任务机制引入的意义所在。其次,与 HCDD 和 ARAE 等同样基于 Hard 机制的多任务模型相比,本文提出的 MTL-SEResNet 模型采用改进的残差块方式,保留了有用原始输入信息,能够自适应的提取故障特征,且降低了模型的复杂度,因而模型的准确率更高、运行时间更短。

3 结 论

本文提出一种基于 CWT 和 MTL-SEResNet 网络的万能式断路器故障程度评估方法。利用 CWT 将振动信号转换成时频图,并引入了多任务学习结构,同时完成故障分类和程度评估任务,得出以下结论。

1) 相较于 STFT、HHT 两种时频分析方法,CWT 时频图的能量分布较为集中,具有更好的时频分辨率,能够凸显时间和频率的细节特征,更适用于万能式断路器振动信号的处理。

2) 所提 MTL-SEResNet 模型基于 Hard 共享机制,实现了故障分类和程度评估任务的互助学习,并通过调整两个任务损失函数的权重比例,达到各任务性能最佳。同时针对模型关键环节 ResNet18 结构的改进,一方面将 Avgpool 改为 GAP 层,在保证准确率的前提下,有效减小了模型的参数量;另一方面,SENet 的引入提升了模型提取有用特征的能力。

3) 与分层模型结构相比,本文模型通过任务之间信息共享,使得模型平均准确率提高了 3% 以上,与其他多任务学习模型相比,改进的网络结构能够自适应的提取故障信息,使得类型及程度准确率分别达到了 99.78% 和 99.36%,且具有更高的运行效率。

考虑到实地采集故障信号的局限性,研究数据均在实验室条件下获得,在今后工作中还应在复杂的应用场合中来验证所提算法的适用性。

参考文献

- [1] LU Y, LI Y. A novel fault diagnosis method for circuit breakers based on optimized affinity propagation clustering[J]. International Journal of Electrical Power & Energy Systems, 2020(118): 105651.
- [2] ZHANG H Y, XIE Y Z, YI T Q, et al. Fault detection for high-voltage circuit breakers based on time-frequency analysis of switching transient E-fields [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(4): 1620-1631.
- [3] GUO W X, WEN F S, LEDWICH G, et al. An analytic model for fault diagnosis in power system considering malfunctions of protective relays and circuit breakers[J]. IEEE Transactions on Power Delivery, 2010, 25(3): 1393-1401.
- [4] ZHANG Y, CHUNG C Y, WEN F S, et al. An analytic model for fault diagnosis in power systems utilizing redundancy and temporal information of alarm messages[J]. IEEE Transactions on Power Systems, 2016, 31(6): 4877-4886.
- [5] 马速良, 武建文, 袁洋, 等. 多振动信息下的高压断路器机械故障随机森林融合诊断方法[J]. 电工技术报, 2020, 35(S2): 421-431.
MA S L, WU J W, YUAN Y, et al. Mechanical fault fusion diagnosis of high voltage circuit breaker using multi-vibration information based on random forest [J]. Transactions of China Electrotechnical Society, 2020, 35(S2): 421-431.
- [6] 杨秋玉, 王栋, 阮江军, 等. 基于振动信号的断路器机械零部件故障程度识别[J]. 电工技术学报, 2021, 36(13): 2880-2892.
YANG Q Y, WANG D, RUAN J J, et al. Fault severity estimation method for mechanical parts in circuit breakers based on vibration analysis [J]. Transactions of China Electrotechnical Society, 2021, 36(13): 2880-2892.
- [7] SHAO K, FU W, TAN J, et al. Coordinated approach fusing time-shift multiscale dispersion entropy and vibrational Harris hawks optimization-based SVM for fault diagnosis of rolling bearing [J]. Measurement, 2020, DOI: 10.1016/j.measurement.2020.108580.
- [8] HU Q, QIN A S, ZHANG Q H, et al. Fault diagnosis based on weighted extreme learning machine with wavelet packet decomposition and KPCA [J]. IEEE Sensors Journal, 2018, 18(20): 8472-8483.
- [9] 孙曙光, 张强, 杜太行, 等. 基于振动信号的低压万能式断路器分合闸故障程度评估方法的研究[J]. 中国电机工程学报, 2017, 37(18): 5473-5482, 5547.
SUN SH G, ZHANG Q, DU T H, et al. Study of evaluation method for low voltage conventional circuit breaker switching fault degree based on vibration signal[J]. Proceedings of the CSEE, 2017, 37(18): 5473-5482, 5547.
- [10] 孔子迁, 邓蕾, 汤宝平, 等. 基于时频融合和注意力机制的深度学习行星齿轮箱故障诊断方法[J]. 仪器仪表学报, 2019, 40(6): 224-230.

- KONG Z Q, DENG L, TANG B P, et al. Fault diagnosis of planetary gearbox based on deep learning with time-frequency fusion and attention mechanism [J]. Chinese Journal of Scientific Instrument, 2019, 40(6): 224-230.
- [11] WEN L, LI X Y, GAO L, et al. A new convolutional neural network-based data-driven fault diagnosis method[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5990-5998.
- [12] 张昌凡, 刘佳峰, 何静, 等. 基于改进卷积双向门控循环网络的轴承故障诊断[J]. 电子测量与仪器学报, 2021, 35(11): 61-67.
- ZHANG CH F, LIU J F, HE J, et al. Improved CNN-BiGRU method for bearing fault diagnosis[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(11): 61-67.
- [13] 王辉, 徐佳文, 严如强. 基于深度宽卷积 Q 网络的行星齿轮箱故障智能诊断方法[J]. 仪器仪表学报, 2022, 43(3): 109-120.
- WANG H, XU J W, YAN R Q. Intelligent fault diagnosis for the planetary gearbox based on the deep wide convolution Q network [J]. Chinese Journal of Scientific Instrument, 2022, 43(3): 109-120.
- [14] 张立智, 徐卫晓, 井陆阳, 等. 基于二维深度卷积网络的旋转机械故障诊断[J]. 机械强度, 2020, 42(5): 1039-1044.
- ZHANG L ZH, XU W X, JING L Y, et al. Rotating machinery fault diagnosis based on two-dimensional convolution neural network [J]. Journal of Mechanical Strength, 2020, 42(5): 1039-1044.
- [15] 赵学智, 叶邦彦, 陈统坚. 短时傅里叶变换的时频聚集性度量准则研究[J]. 振动. 测试与诊断, 2017, 37(5): 948-956, 1065.
- ZHAO X ZH, YE B Y, CHEN T J. Study on measure rule of time-frequency concentration of short yime Fourier transform [J]. Journal of Vibration, Measurement & Diagnosis, 2017, 37(5): 948-956, 1065.
- [16] 张龙, 甄灿壮, 熊国良, 等. 基于深度时频特征的机车轴承故障诊断[J]. 交通运输工程学报, 2021, 21(6): 247-258.
- ZHANG L, ZHEN C ZH, XIONG G L, et al. Locomotive bearing fault diagnosis based on deep time-frequency characteristics [J]. Journal of Traffic and Transportation Engineering, 2021, 21(6): 247-258.
- [17] 柳絮, 王坚, 李文. 集成变分模态分解和希尔伯特-黄变换的结构振动时频提取模型[J]. 武汉大学学报(信息科学版), 2021, 46(11): 1686-1692.
- LIU X, WANG J, LI W. A time frequency extraction model of structural vibration combining VMD and HHT[J]. Geomatics and Information Science of Wuhan University, 2021, 46(11): 1686-1692.
- [18] 康守强, 刘哲, 王玉静, 等. 基于改进 DQN 网络的滚动轴承故障诊断方法[J]. 仪器仪表学报, 2021, 42(3): 201-212.
- KANG SH Q, LIU ZH, WANG Y J, et al. A fault diagnosis method of rolling bearing based on the improved DQN network [J]. Chinese Journal of Scientific Instrument, 2021, 42(3): 201-212.
- [19] GOU L F, LI H H, ZHENG H, et al. Aeroengine control system sensor fault diagnosis based on CWT and CNN[J]. Mathematical Problems in Engineering, 2020, DOI: 10.1155/2020/5357146.
- [20] YANG Q Y, RUAN J, ZHUANG Z, et al. Condition evaluation for opening damper of spring operated high-voltage circuit breaker using vibration time-frequency image [J]. IEEE Sensors Journal, 2019, 19(18): 8116-8126.
- [21] ZHAO M, ZHONG S, FU X, et al. Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis[J]. IEEE Transactions on Industrial Electronics, 2021, 68(3): 2587-2597.
- [22] 田科位, 董绍江, 姜保军, 等. 基于改进深度残差网络的轴承故障诊断方法[J]. 振动与冲击, 2021, 40(20): 247-254.
- TIAN K W, DONG SH J, JIANG B J, et al. A bearing fault diagnosis method based on an improved depth residual network [J]. Journal of Vibration and Shock, 2021, 40(20): 247-254.
- [23] GUO X J, CHEN L, SHEN C. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis [J]. Measurement, 2016, 93: 490-502.
- [24] GAN M, WANG C, ZHU C AN, et al. Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings [J]. Mechanical Systems and Signal Processing, 2016, 72-73: 92-104.
- [25] CARUANA R. Multitask learning [J]. Machine Learning, 1997, 28(1): 41-75.
- [26] WANG H, LIU Z, PENG D, et al. Feature-level attention-guided multitask CNN for fault diagnosis and

working conditions identification of rolling bearing [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, DOI: 10.1109/TNNLS.2021.3060494.

- [27] WEN L, LI X, GAO L. A new two-level hierarchical diagnosis network based on convolutional neural network [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(2): 330-338.
- [28] HE K, ZHANG S, REN S, et al. Deep residual learning for image recognition [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
- [29] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [30] HUANG N, FANG L, CAI G, et al. Mechanical fault diagnosis of high voltage circuit breakers with unknown fault type using hybrid classifier based on LMD and time segmentation energy entropy [J]. Entropy, 2016, 18(9), DOI: 10.3390/e18090322.
- [31] KONG X, LI X, ZHOU Q, et al. Attention recurrent autoencoder hybrid model for early fault diagnosis of rotating machinery [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-10.

作者简介



孙曙光, 分别在 2002 年、2005 年和 2009 年于河北工业大学获得学士学位、硕士学位和博士学位, 现为河北工业大学副教授, 主要研究方向为智能检测与故障诊断。
E-mail: sunshuguang_2000@163.com

Sun Shuguang received his B. Sc. degree, M. Sc. degree, and Ph. D. degree all from Hebei University of Technology in 2002, 2005, and 2009, respectively. He is currently an associate professor at Hebei University of Technology. His main research interests include intelligent detection and fault diagnosis.



张婷婷, 2019 年于防灾科技学院获得学士学位, 现为河北工业大学硕士研究生, 主要研究方向为检测技术与自动化装置。

E-mail: 786102517@qq.com

Zhang Tingting received her B. Sc. degree from the Institute of Disaster Prevention in 2019. She is currently a M. Sc. candidate at Hebei University of Technology. Her main research interests include measurement technology and automatic equipment.



王景芹 (通信作者), 分别在 1984 年、1986 年和 1997 年于河北工业大学获得学士学位、硕士学位和博士学位, 现为河北工业大学教授, 主要研究方向为电接触和可靠性。

E-mail: jqwang@hebut.edu.cn

Wang Jingqin (Corresponding author) received her B. Sc. degree, M. Sc. degree, and Ph. D. degree all from Hebei University of Technology in 1984, 1986, and 1997, respectively. She is currently a professor at Hebei University of Technology. Her main research interests include electrical contact and reliability.