

基于 ECAProNet 模型的小样本变工况 轴承故障诊断

王 勉, 吴东升, 王笑臣

(沈阳理工大学 自动化与电气工程学院, 沈阳 110159)

摘要: 针对小样本变工况轴承故障诊断中易出现过拟合的问题, 提出一种在原型网络(ProNet)中加入高效通道注意力机制(ECA)的轴承故障诊断方法(ECAProNet)。基于度量学习的思想, 在改进原型网络框架下将通过特征提取模块的样本信号映射至特征度量空间, 在该空间内确定查询样本与各类原型间的欧式距离, 得到损失函数, 进而优化特征提取网络框架; 为达到元学习目的, 采用基于 episodes 的训练策略, 将算法泛化到不同工况的测试诊断中。在 CW-RU 数据集上设置 5-way 5-shot 和 5-way 1-shot 验证实验, 结果表明, ECAProNet 在小样本变工况轴承故障诊断中表现出较好的性能。

关键词: 故障诊断; 小样本学习; 变工况; 原型网络; 高效通道注意力机制

中图分类号: TP277.3 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2024.01.004

Bearing Fault Diagnosis under Variable Working Conditions with Few Samples Based on ECAProNet Model

WANG Mian, WU Dongsheng, WANG Xiaochen

(Shenyang Ligong University, Shenyang 110159, China)

Abstract: For the problem of over-fitting in bearing fault diagnosis under variable working conditions in the case of few samples, a bearing fault diagnosis method (ECAProNet) with efficient channel attention mechanism (ECA) added to the prototype network (ProNet) is proposed. Based on the idea of metric learning, the sample signal of the feature extraction module is mapped to the feature metric space under the framework of the improved prototype network. In this space, the Euclidean distance between the query sample and various prototypes is determined, and the loss function is obtained to optimize the feature extraction network framework. In order to achieve the purpose of meta-learning, the training strategy based on episodes is used to generalize the algorithm to the test diagnosis of different working conditions. The 5-way 5-shot and 5-way 1-shot verification experiments are set on the CWRU dataset. The results show that ECAProNet has good performance in bearing fault diagnosis with few samples and variable working conditions.

Key words: fault diagnosis; few-shot learning; variable working conditions; prototypical network; efficient channel attention

滚动轴承是广泛应用于各类机械设备中的重要基础部件, 在设备的稳定可靠运行中发挥重要

作用。滚动轴承在使用过程中易发生碎片剥落、磨损产生裂纹及凹坑,造成机械设备无法正常运行^[1],甚至引起安全事故。因此,对滚动轴承进行故障诊断具有重要的实际意义。

传统机器学习故障诊断方法只适用于某一类机械,更换到另一类机械后不再有良好的诊断能力^[2-3]。近些年,深度学习算法被广泛应用于轴承的故障诊断^[4-6],其可通过深度神经网络(deep neural network, DNN)自动完成特征提取,但深度学习具有十分严重的数据依赖性,数据不足会导致训练过程中发生过拟合现象,而实际生产中往往难以提供大量数据来满足深度学习的要求。由此,人们提出了小样本分类方法。

Snell 等^[7]于 2017 年首次提出基于度量的原型网络,并将其用于图像的少样本分类。该网络将带标签样本与待识别样本映射至同一空间,将同类样本均值作为类原型,通过衡量待识别样本与各类原型间的欧氏距离进行分类。由于其泛化能力出色且易与其他模型相结合,诸多学者将其应用于少样本轴承故障诊断中。Chen 等^[8]提出将轴承振动信号通过快速傅里叶变换(FFT)转变为频域信号,使用马哈拉诺比斯距离代替欧氏距离,提高了诊断准确率。乔春阳^[9]提出基于金字塔分割注意力的半监督原型网络,使用未标记样本对初始原型进行调整,减少了构建网络模型时对标记样本的依赖。华超等^[10]提出基于密度加权的原型网络,对每个样本赋予不同的权值,对代表性大的样本赋予更大权值,对代表性小的样本赋予更小的权值,以此缓解样本代表性不足的问题。Wang 等^[11]提出原型孪生网络,在孪生网络中增加了一个原型模块,以学习每个类的高质量原型表示,在少样本学习实验中表现出较好的性能。

上述方法均在小样本学习领域取得了较好的效果,但文献[8]提出的网络模型需要构建一个卷积神经网络(CNN)作为特征提取网络,再应用另一 CNN 作为方差预测网络,一定程度上增加了模型的复杂程度,加大了训练的工作量。文献[9]提出的模型中每个类别需要 3~10 个标记样本,以满足训练阶段的要求,无法应对特殊情况下仅有某类别且仅存在唯一样本的变工况轴承故障诊断。文献[10]提出的改进原型网络在进行 50 次迭代后,准确率才有明显提升,时间成本较高。文献[11]提出的网络模型在 1-shot 情况下的分类准确率过低,仅为 48.7%。

针对上述问题,本文在原型网络(ProNet)的基础上,提出在预学习阶段引入高效通道注意力机制(ECA)^[12],对样本的特征进行充分挖掘,增强对具有代表性特征的关注度,提高模型分类准确率,同时不增加模型的复杂程度。本文提出的故障诊断方法简称为 ECAProNet。

1 原型网络

1.1 数据集的划分

ProNet 采用元学习的训练方法进行小样本训练,可通过极少数的样本确定类原型,通过衡量待识别样本与类原型间的欧氏距离,采用近邻策略进行识别分类,其分类原理如图 1 所示。ProNet 算法采用基于 episodes 的训练策略,将数据集划分为元训练集 D_{train} 和元测试集 D_{test} ,且训练集与测试集样本空间不相交,以使训练过程中优化好的网络模型能够泛化到测试环境。

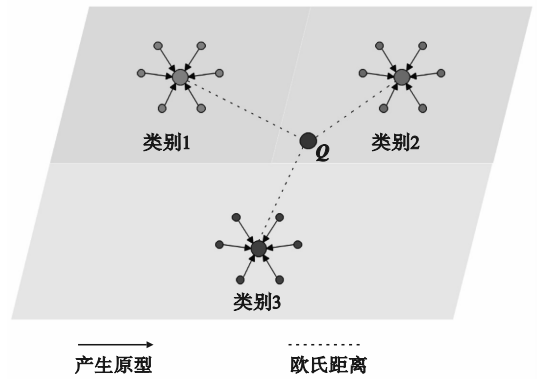


图 1 原型网络分类原理图

Fig. 1 Prototypical network classification schematic diagram

元训练数据集 D_{train} 分为支持集 S 和查询集 Q ,支持集表示为 $S = \{x_i, y_i\}_{i=1}^{n_s}$,查询集表示为 $Q = \{\tilde{x}_i, \tilde{y}_i\}_{i=1}^{n_q}$ 。其中: x_i 和 \tilde{x}_i 分别为支持集和查询集样本, $x_i \in \mathbf{R}^D$, $\tilde{x}_i \in \mathbf{R}^D$, D 为样本特征向量的维数; y_i 和 \tilde{y}_i 分别表示 x_i 和 \tilde{x}_i 对应类别的标签; n_s 为每类支持样本数; n_q 为每类查询样本数。若在支持集 S 中包含了 N 个类别且每个类别中包含了 K 个样本,则将此问题称为 N -way K -shot 问题。

1.2 原型计算与分类

支持集经过特征提取网络后, N 个类别中每一个样本从 D 维空间映射到 M 维空间(M 为原型特征向量的维数),并得到特征向量,在 M 维空间中通过计算支持集样本特征向量的平均值得到原

型。原型计算式为

$$L_c = \frac{1}{K} \sum_{i=1}^K f_\phi(x_i) \quad (1)$$

式中: L_c 表示第 c 类的原型, $L_c \in \mathbf{R}^M$; K 为支持集中 c 类别中抽取的样本数量; $f_\phi(\cdot)$ 表示具有可学习参数 ϕ 的嵌入函数。

计算出查询集中样本与各类原型间的欧氏距离, 利用 softmax 分类器得到查询集样本 \tilde{x}_i 属于第 c 类别的概率 P_ϕ , 具体计算式为

$$P_\phi(y = c | \tilde{x}_i) = \frac{\exp[-d(f_\phi(\tilde{x}_i), L_c)]}{\sum_{c'=1}^N \exp[-d(f_\phi(\tilde{x}_i), L_{c'})]} \quad (2)$$

式中 $d(\cdot, \cdot)$ 为距离函数。

1.3 损失函数的计算

为使分类概率最大化, 利用损失函数使网络不断迭代更新自身参数, 取得最小损失, 损失函数 J 计算式为

$$J = -\log P_\phi(y = c | \tilde{x}_i) \quad (3)$$

2 ECAProNet 算法

2.1 ECANet 算法

ECA 使用不降维的局部跨通道交互策略, 解决了降维造成的通道注意力学习能力差的问题。该方法涉及参数少且效果显著, 一定程度上降低了模型的复杂程度。

ECA 网络结构如图 2 所示。图中: χ 为输入特征; $\tilde{\chi}$ 为带有注意力的特征; W 、 H 、 C 分别为卷积块的宽度、高度及通道维数。ECA 模型使用全局平均池化(GAP)不降维聚合卷积特征, 确定卷积核大小 q , 进行一维卷积, 通过 Sigmoid 函数学习通道注意力, 得到通道的权重, 处理通道权重与输入特征后即可得到带注意力的特征。

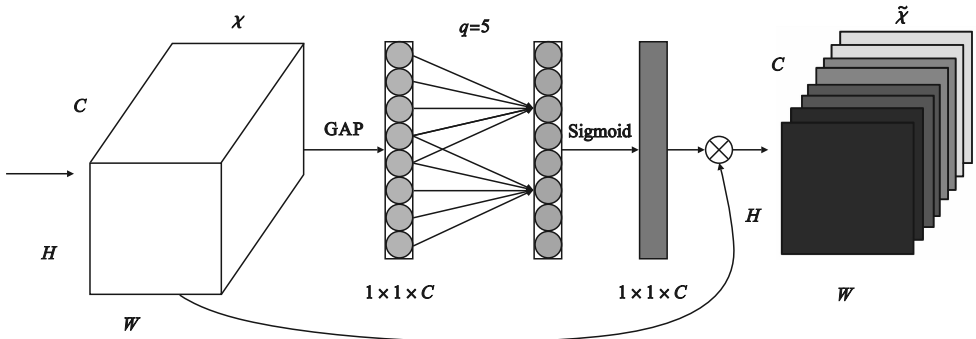


图2 ECA 网络结构

Fig.2 ECA network structure

ECA 通过大小为 q (其大小决定局部跨通道交互的覆盖范围) 的卷积核快速卷积, 实现跨通道交互。 q 与通道维数 C 成正比, 映射关系为

$$C = \varphi(q) = 2^{(\gamma \times q - b)} \quad (4)$$

$$q = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (5)$$

式中: $\lfloor t \rfloor_{\text{odd}}$ 表示距离 t 最近的奇数; γ 和 b 为自定义参数, 本文分别设置为 2 和 1。

2.2 ECAProNet 算法

本文采用 ResNet^[13] 作为特征提取网络, 将 ECA 引入 ResNet 网络中。通过特征提取网络 (ECA-ResNet) 将样本从 D 维空间映射至 M 维空间, 依据式(1)计算出类原型, 通过衡量查询集样本与类原型之间的欧氏距离, 采用 softmax 函数计算查询集样本属于各类原型的概率, 计算出查询集样本的负对数概率, 得到损失函数, 将损失函数

反馈到特征提取网络中, 优化自身参数, 达到网络的迭代更新。ECAProNet 算法框图如图 3 所示。

ResNet 引入残差块并采用跳跃连接将输入数据直接传递到后续层, 增强信息流动性, 提高训练精度。基于残差网络的特性^[13], 在残差块中插入 ECA 模块, 在不增加模型复杂度的前提下, 能够在少量数据中提取到更具代表性的特征; 将带有 ECA 模块的 ResNet 作为原型网络的特征提取网络, 在有限样本数据中提取更有用的滚动轴承特征信息。

2.3 基于 ECAProNet 的变工况诊断流程

本文提出的基于 ECAProNet 的诊断流程如图 4 所示, 具体步骤如下。

步骤 1 将采集到的一维振动信号转变为灰度图像, 以便采集到更多有用信息。

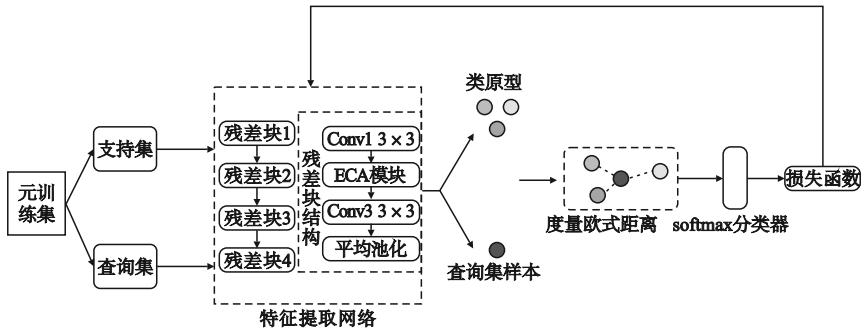


图 3 ECAProNet 算法框图

Fig. 3 ECAProNet algorithm block diagram

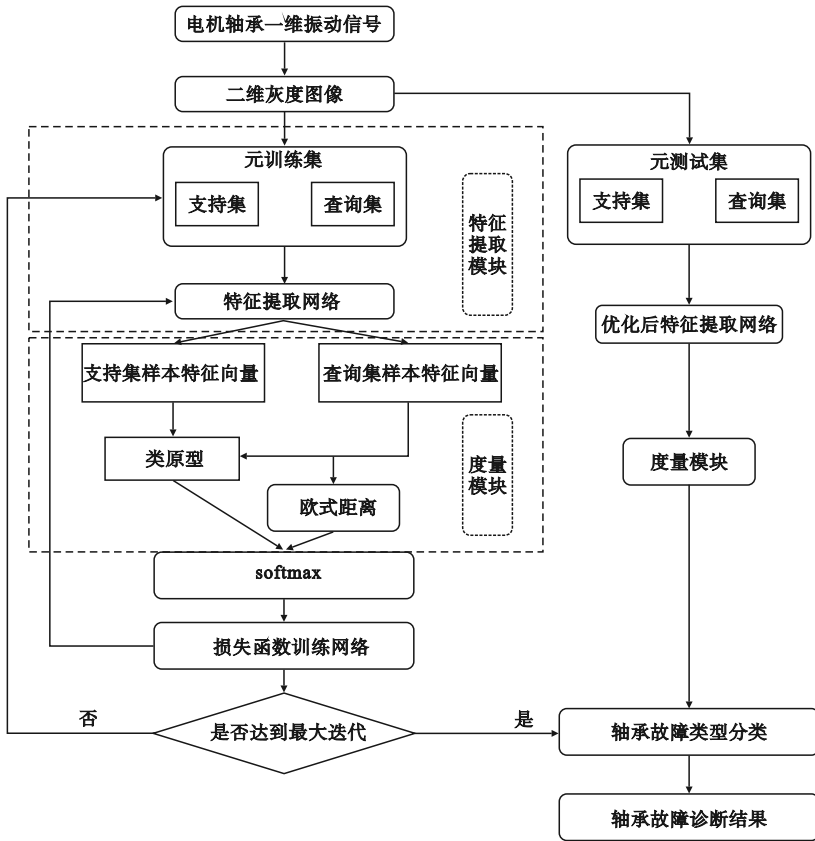


图 4 ECAProNet 故障诊断流程框图

Fig. 4 Block diagram of ECAProNet fault diagnosis process

步骤 2 在预训练阶段,于采样频率 12 kHz 下进行采样,共分为 4 种工况,每种工况包含 5 个故障类型,在 20 类故障类型中随机选取 N 类样本,并从 N 类样本中随机抽取 K 个样本构成支持集 S ,再从每类样本中随机抽取样本构成查询集 Q ,作为特征提取网络的输入。

步骤 3 由 ECA-ResNet 特征提取网络将支持集 S 和查询集 Q 中的样本映射到统一特征度量空间,得到支持集 S 与查询集 Q 中各样本的特征向量。

步骤 4 由式(1)计算出支持集 S 中每类样本的平均向量作为类原型,计算查询集 Q 中各样

本的特征向量与类原型间的欧氏距离后输入 softmax 函数;由式(2)得到查询集中样本属于每类原型的概率;通过式(3)计算得到损失函数并反馈到特征提取网络,对其进行训练,完成模型的迭代更新,优化自身参数。

步骤 5 在测试阶段,经基于 episodes 训练策略优化好的特征提取网络将 D_{test} 中含有带标签信号的支持集和查询集作为输入,映射到度量空间,确定查询集中无标签样本的调制样式,确定轴承故障分类,得到变工况故障诊断结果。

3 实验验证

3.1 环境配置

本文实验在 Pytorch 框架下进行搭建,硬件设施为 Intel Core i7-8700CPU、NVIDIA GTX1080ti 显卡的计算机。

3.2 数据集

本实验采用的数据集为凯斯西储大学 (CRWU) 的开放轴承数据库^[14],选用 12 kHz 采样频率下的轴承振动信号作为训练数据集,选用 48 kHz 采样频率下的轴承振动信号作为测试数据集,验证模型在不同工况下的故障诊断效果。设置 4 种工况 (A、B、C、D) 及 5 种轴承故障类型,如表 1 和表 2 所示。

表 1 CRWU 轴承的 4 种工况

Table 1 Four kinds of working conditions of CRWU bearings

工况	负载/W	转速/(r·min ⁻¹)
A	0	1 797
B	735	1 772
C	1 470	1 750
D	2 205	1 730

表 2 CRWU 的故障类型

Table 2 Fault types of CRWU

故障代号	故障类型
RB	滚动体故障
IR	内圈故障
OR@3	外圈故障(正交方向)
OR@6	外圈故障(中心方向)
OR@12	外圈故障(相对方向)

3.3 模型对比实验及结果分析

将 ECAProNet 与其他四种小样本诊断方法进行对比,分别为 ProNet、与模型无关的元学习模型 (MAML)^[15]、基准模型 (Baseline)、多任务学习模型 (MTL)^[16],对比实验结果如表 3 和表 4 所示。为得到较为稳定、客观的诊断准确率,将 5-way 5-shot 实验与 5-way 1-shot 实验分别在训练阶段进行 15 次迭代,在测试阶段进行 10 轮测试,取平均准确率作为实验结果。对比表 3 和表 4 可以看出,在 5-way 5-shot 样本抽取下的变工况轴承故障诊断准确率绝大多数高于 5-way 1-shot。原型网络中支持集样本数据的增加可为特征提取网络提供更多的特征信息,提高诊断的准确率,5-way 1-shot 的实验结果恰恰是样本数据不足所造

成的过拟合现象。

表 3 5-way 5-shot 变工况轴承故障诊断结果

Table 3 5-way 5-shot variable working conditions bearing fault diagnosis results

任务	准确率/%				
	Baseline	MTL	MAML	ProNet	ECAProNet
A→B	80.83	88.00	85.60	82.76	88.52
A→C	85.41	91.66	90.93	83.05	90.04
A→D	76.33	92.50	90.53	81.56	86.36
B→A	83.21	77.59	89.06	86.00	90.00
B→C	88.06	75.44	92.53	85.60	89.36
B→D	86.50	90.13	87.60	87.36	91.60
C→A	76.48	77.52	75.61	72.96	81.92
C→B	89.13	88.00	88.38	85.50	92.04
C→D	91.50	90.00	95.20	84.80	92.88
D→A	80.40	80.67	77.12	74.28	82.68
D→B	86.44	88.23	88.24	83.40	86.88
D→C	82.57	80.15	91.28	82.20	84.24

表 4 5-way 1-shot 变工况轴承故障诊断结果

Table 4 5-way 1-shot variable working conditions bearing fault diagnosis results

任务	准确率/%				
	Baseline	MTL	MAML	ProNet	ECAProNet
A→B	78.58	85.33	83.54	79.64	86.80
A→C	65.10	78.19	83.01	72.32	79.20
A→D	82.88	80.24	82.27	73.52	84.68
B→A	84.26	75.33	82.53	74.80	80.00
B→C	83.49	89.65	87.05	78.84	82.41
B→D	85.05	87.34	79.88	78.74	81.92
C→A	74.65	85.77	71.92	68.80	74.40
C→B	82.57	87.27	88.13	84.72	88.40
C→D	85.20	90.14	91.71	84.68	85.60
D→A	74.15	72.16	75.51	73.64	77.04
D→B	86.96	89.77	83.73	80.40	86.60
D→C	85.08	84.33	93.21	77.10	79.08

与 ProNet 模型相比,加入注意力机制的 ECAProNet 在诊断性能方面有着明显提升;与其他三种模型相比,ECAProNet 的诊断性能亦表现较佳。ECA 模块的加入使改进后的特征提取网络能够关注到更多有用信息,在提升性能的同时并不增加模型的复杂程度,一定程度上保证了整体模型不会因为复杂度过高造成过拟合现象。

观察 5-way 1-shot 实验结果可以发现,ProNet 模型的诊断准确率低于其他四种模型,这是因为

每类样本数据只有 1 组,导致度量空间中样本特征向量的平均值即是抽样样本的特征向量,某工况下训练好的故障分类模型无法准确对不同工况下测试样本进行故障分类,在加入 ECA 模块后,诊断准确率则有明显的提升。

4 结论

针对小样本环境下的变工况轴承故障诊断,基于元学习和度量学习的思想提出了 ECAProNet 诊断算法。通过加入 ECA 模块提升网络模型的特征提取能力,得出更具代表性的类原型,提升诊断准确率。在 CWRU 轴承数据集上的实验结果表明:在小样本环境下,网络模型存在过拟合问题;相较于原始的 ProNet 算法,本文提出的 ECA-ProNet 算法故障诊断准确率提升较为明显,同时提高了模型在训练时的泛化能力。

参考文献 (References):

[1] YU C C, NING Y Q, QIN Y, et al. Multi-label fault diagnosis of rolling bearing based on meta-learning[J]. *Neural Computing and Applications*, 2021, 33(10): 5393 – 5407.

[2] 李军宁, 罗文广, 陈武阁. 面向振动信号的滚动轴承故障诊断算法综述[J]. *西安工业大学学报*, 2022, 42(2): 105 – 122.

LI J N, LUO W G, CHEN W G. Overview of algorithms for rolling bearing fault diagnosis based on vibration signal[J]. *Journal of Xi'an Technological University*, 2022, 42(2): 105 – 122. (in Chinese)

[3] 柳秀, 马善涛, 谢怡宁, 等. 面向轴承故障诊断的深度学习方法[J]. *哈尔滨理工大学学报*, 2022, 27(4): 118 – 124.

LIU X, MA S T, XIE Y N, et al. Deep learning method for bearing fault diagnosis[J]. *Journal of Harbin University of Science and Technology*, 2022, 27(4): 118 – 124. (in Chinese)

[4] 黄扣, 袁伟, 陈红卫. 基于深度学习的轴承故障智能诊断方法研究[J]. *计算机与数字工程*, 2022, 50(8): 1827 – 1832.

HUANG K, YUAN W, CHEN H W. Research on bearing fault intelligent diagnosis method based on deep learning[J]. *Computer & Digital Engineering*, 2022, 50(8): 1827 – 1832. (in Chinese)

[5] 魏永合, 宫俊宇. 基于 CNN-LSTM-Attention 的滚动轴承故障诊断[J]. *沈阳理工大学学报*, 2022, 41(4): 73 – 77.

WEI Y H, GONG J Y. Fault diagnosis in rolling bearing based on CNN-LSTM-Attention[J]. *Journal of Shenyang Ligong University*, 2022, 41(4): 73 – 77. (in Chinese)

[6] 孙波, 姜宇宏, 杨青. SSA 优化融合 CNN 的电机轴承故障诊断方法研究[J]. *沈阳理工大学学报*, 2021, 40(3): 9 – 13.

SUN B, JIANG Y H, YANG Q. Research on fault diagnosis method of motor bearing based on SSA optimization and CNN[J]. *Journal of Shenyang Ligong University*, 2021, 40(3): 9 – 13. (in Chinese)

[7] SNELL J, SWERSKY K, ZEMEL R S. Prototypical networks for few-shot learning[C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, CA, USA; MIT Press, 2017: 4080 – 4090.

[8] CHEN Y B, HONG Y, LONG J Y, et al. Few shot learning for novel fault diagnosis with a improved prototypical network[C]//*2021 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence(IC-SMD)*. Nanjing, China; IEEE, 2021: 1 – 6.

[9] 乔春阳. 基于小样本学习的不同型号滚动轴承故障诊断方法研究[D]. 哈尔滨: 哈尔滨理工大学, 2022.

[10] 华超, 刘向阳. 基于密度加权原型网络的小样本学习算法[J]. *计算机技术与发展*, 2022, 32(9): 8 – 13.

HUA C, LIU X Y. Few-shot learning based on density-weighted prototypical network[J]. *Computer Technology and Development*, 2022, 32(9): 8 – 13. (in Chinese)

[11] WANG J H, ZHAI Y P. Prototypical Siamese networks for few-shot learning[C]//*2020 IEEE 10th International Conference on Electronics Information and Emergency Communication(ICEIEC)*. Beijing, China; IEEE, 2020: 178 – 181.

[12] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//*2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR)*. Seattle, WA, USA; IEEE, 2020: 11531 – 11539.

[13] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*. Las Vegas, NV, USA; IEEE, 2016: 770 – 778.

[14] Case Western Reserve University. Bearing Data Center[DB/OL]. [2023-03-15]. <https://engineering.case.edu/bearingdatacenter/download-data-file>.

[15] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[C]//*Proceedings of the 34th International Conference on Machine Learning*. New York, USA; ACM, 2017: 1126 – 1135.

[16] LI W H, BILEN H. Knowledge distillation for multi-task learning[C]//*Computer Vision; ECCV 2020 Workshops*. Glasgow, UK; Springer, 2020: 163 – 176.

(责任编辑: 宋颖韬)