

小样本下 SE-ResNet 与元迁移学习的 变工况轴承故障诊断

刘 臻, 彭珍瑞, 王圣杰

(兰州交通大学机电工程学院, 甘肃 兰州 730070)

摘要: 针对轴承在变工况下样本分布不同、故障样本少和一些小样本算法特征提取有限, 导致轴承故障诊断精度低及模型泛化能力弱的问题, 提出了小样本下嵌入压缩、激励的残差网络 (SE-ResNet) 与元迁移学习 (MTL) 的变工况轴承故障诊断方法。将采集的不同工况下轴承一维振动信号通过连续小波变换 (CWT) 转换成对应工况下的时频图像, 从而将轴承故障诊断问题转换为图像识别问题; 引入压缩-激励注意力机制, 构建了一种 SE-ResNet 的骨干网络模型, 以聚焦于更有效的特征通道, 增强特征提取表征能力; 借助迁移学习能提供良好的深层网络初始参数和元学习能快速学习的优势, 依次进行预训练与元迁移训练, 得到利用少量样本微调便能达到高精度的元迁移网络, 进而实现变工况下轴承的故障诊断; 通过两个基准数据集和实验室搭建的轴承故障模拟试验台进行验证, 并与其他方法进行对比分析, 结果表明, 所提方法在小样本、变工况下对轴承故障诊断具有更高的识别精度和泛化性能。

关键词: 轴承故障诊断; 连续小波变换; 元迁移学习; 变工况; 小样本

中图分类号: TH133.3 **文献标志码:** A **文章编号:** 1004-4523(2025)06-1199-13

DOI: 10.16385/j.cnki.issn.1004-4523.2025.06.008

Bearing fault diagnosis under few-shot and variable working conditions using SE-ResNet and Meta-Transfer learning

LIU Zhen, PENG Zhenrui, WANG Shengjie

(School of Mechanical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: Traditional bearing fault diagnosis methods often suffer from low accuracy and weak model generalization under varying working conditions due to diverse sample distributions, scarcity of fault samples, and limited feature extraction capabilities of some few-shot learning algorithms. To address these challenges, this paper proposes a novel method for variable condition bearing fault diagnosis that combines a squeeze-and-excitation residual network (SE-ResNet) with meta-transfer learning (MTL). One-dimensional bearing vibration signals collected under different working conditions are converted into time-frequency images using continuous wavelet transform (CWT), thereby transforming the bearing fault diagnosis task into an image recognition problem. A squeeze-and-excitation (SE) attention mechanism is introduced to construct an SE-ResNet backbone network model. This focuses on more effective feature channels, thereby enhancing feature extraction and representation capabilities. Leveraging the advantages of transfer learning (which provides robust initial deep network parameters) and meta-learning (which enables rapid adaptation), the model undergoes sequential pre-training and meta-transfer training. This process yields a high-precision meta-transfer network that can be fine-tuned with only a small number of samples, ultimately achieving accurate bearing fault diagnosis under variable working conditions. The proposed method is validated using two benchmark datasets and a bearing fault simulation test bench developed in the laboratory. Comparative analysis with other methods demonstrates that the proposed method exhibits higher recognition accuracy and superior generalization performance for bearing fault diagnosis under both few-shot and variable working conditions.

Keywords: bearing fault diagnosis; continuous wavelet transform; meta-transfer learning; variable working conditions; few-shot

现代旋转机械设备向着信息化、智能化的方向前进。滚动轴承作为旋转机械中的重要零部件, 已被广泛应用于航空航天、能源化工和交通运输等行

业^[1]。因其运行环境恶劣, 极易发生故障, 可能会导致巨大的经济损失, 甚至人员伤亡。另外, 在实际运行环境中, 转速以及载荷等运行条件不断变化, 故轴

收稿日期: 2024-12-12; 修订日期: 2025-03-03

基金项目: 甘肃省自然科学基金资助项目(20JR10RA209); 甘肃省科技厅优秀博士生项目(23JRRA890, 25JRRA215)

承常在变工况下运行^[2-3]。因此,开展变工况下轴承故障诊断研究具有重要理论研究价值和工程应用意义。

传统的基于机器学习的故障诊断是通过经验模态分解、变分模态分解和奇异值分解等方法对振动信号进行处理,提取敏感故障特征并输入到支持向量机、随机森林、人工神经网络等模型中进行训练,从而实现轴承故障诊断^[4]。此类方法存在两大局限性:其一,故障特征的提取严重依赖于专家经验;其二,特征提取和模型优化相互独立,不能联合优化,这限制了智能故障诊断在实际工程中的应用。深度学习克服了传统机器学习的缺陷,因其强大的特征提取能力已成功应用于轴承故障诊断。赵敬娇等^[5]将一维原始振动信号输入到具有残差连接的多个一维卷积层自动提取特征,从而实现了轴承故障诊断。但原始振动信号并没有考虑到时间相关性,且研究表明二维图像具有更高的识别精度^[6]。康守强等^[7]利用短时傅里叶变换(short-time Fourier transform, STFT)将轴承原始振动信号转换为二维时频图像,由残差网络(ResNet)提取故障特征。但是STFT具有窗口固定的特性,使其分析非平稳信号具有一定的局限性^[8]。连续小波变换(continuous wavelet transform, CWT)具有多分辨率特性,能够在时频域上提供具有表征性的特征信息。XU等^[9]通过CWT将轴承振动信号转换为时频图像,结合卷积神经网络(convolutional neural network, CNN)和深度森林模型进行特征提取和分类识别。CHENG等^[10]通过CWT将振动信号转换为时频图像,利用局部二元卷积神经网络完成了自适应的故障诊断。以上方法均建立在大量的训练样本和恒定工况下,而在实际运行环境中,样本数据十分稀缺且工作条件不断变化,这严重影响了故障诊断算法的性能。

迁移学习可利用大量标注数据的基类预训练一个深度网络模型,并应用于新环境。安文杰等^[11]利用已知风电组数据预训练多尺度卷积神经网络,并将其迁移至目标风电机组以实现故障诊断。一些学者预训练ResNet^[12-13],通过使用少量目标样本对目标域进行微调,获得了良好的性能。基于领域适应的方法是迁移学习的主流方法之一,其基本思想是通过学习源领域和目标领域之间的差异来实现模型的迁移。该方法在跨工况的故障诊断中取得了较好的效果,减少了对目标域标记样本数量的依赖,但需要大量源域数据来学习领域不变特征且对每个新环境都需重新对齐特征。在样本极少或两域分布差异较大的情况下,该算法容易出现负迁移^[14]。近年来,受人类学习的启发而提出了一种小样本学习方法——元学习,旨在使模型学会学习,通过少量的训练样本

设计能够快速学习新技能以适应新环境的模型。相较基于领域适应的方法,元学习在处理小样本和变工况时展现出显著优势,具有更好的泛化性、适应性和鲁棒性^[15]。经典模型是由FINN等^[16]提出的模型无关的元学习(model-agnostic meta-learning, MAML)方法,其目标是寻找对任务变化敏感的初始模型参数,实现仅用少量样本就能快速适应新任务的目的。CHEN等^[17]提出了基于MAML的电机轴承故障诊断方法,以4层的CNN为骨干网络,用于识别不同工况下电机轴承故障。SU等^[18]对采集的轴承振动信号进行了重建,并采用递归元学习方法实现了变工况下小样本的智能故障诊断。HU等^[19]提出了一种任务排序元学习的小样本故障诊断方法,将元训练任务从易到难进行排序,通过不断增加任务难度以获得更好的知识适应性。尽管以上基于元学习的方法取得了较好的效果,但是为避免过拟合,每个任务都是由低复杂度的基础学习器建模,从而无法使用更深入、更强大的架构^[15]。

上述研究虽在轴承故障诊断中取得了较好的效果,但仍存在以下不足:(1)单独地从元学习或迁移学习层面分析并解决了变工况下轴承故障诊断问题,但没有考虑到二者各自的优势;(2)算法模型提取的特征有限,利用已有工况数据识别其余工况故障的泛化能力较弱。

基于以上分析,本文参考文献[20],利用CWT将轴承原始信号转换为二维时频图像,引入压缩-激励(squeeze-excitation, SE)注意力模块与残差网络作为骨干网络,结合元学习和迁移学习的各自优势,利用已知工况数据对网络模型进行预训练,以提供良好的初始化参数,并将已训练的特征提取器进行参数迁移,以元学习方式对小样本任务训练,得到具有较强泛化能力的元迁移网络模型,进而实现变工况下的轴承故障诊断。通过基准数据集和实验室轴承故障模拟试验台数据验证了所提方法的有效性和优越性。

1 基础理论

1.1 连续小波变换

连续小波变换是基于小波变换的信号处理技术,适用于处理非平稳、非线性信号。其具有时频局部化和多分辨率特性,克服了傅里叶变换只有频域没有时域的缺点,以及短时傅里叶变换不能在时域和频域上局域化的缺陷,能更好地分析轴承运行过程中的时变信号。CWT的原理是通过小波基函数 $\psi(t)$ 对原始振动信号 $f(t)$ 进行卷积运算提取一系列的小

波系数 W_f , 从而构建一个具有良好表征性的时频信号^[21], 其表达式为:

$$W_f(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \overline{\psi_{a,b}\left(\frac{t-b}{a}\right)} dt \quad (1)$$

式中, t 为时间; a 为尺度因子; b 为平移因子; $\overline{\psi_{a,b}(t)}$ 为小波基函数 $\psi_{a,b}(t)$ 的复共轭。

小波基函数应尽量与轴承故障时产生的振动脉冲相似, Morlet 小波适合刻画信号的快速振荡特性, 故选其作为小波基函数, 表达式为:

$$\psi(t) = \pi^{-\frac{1}{4}} \exp(i\omega_0 t) \exp\left(-\frac{t^2}{2}\right) \quad (2)$$

式中, ω_0 为小波的中心频率。

1.2 残差网络

CNN 已被广泛应用于计算机视觉、图像识别等领域。深度卷积神经网络得到了蓬勃发展, 随着网络深度的加深, 训练模型时会出现因梯度消失或梯度爆炸而导致网络退化的问题。残差网络有效地解决了该问题, 其核心思想是将网络进行模块划分, 并引入了跳跃连接, 可将输入模块的信息直接绕路输出, 保护了信息的完整性, 该模块被称为残差模块^[22-23], 其结构如图 1 所示。

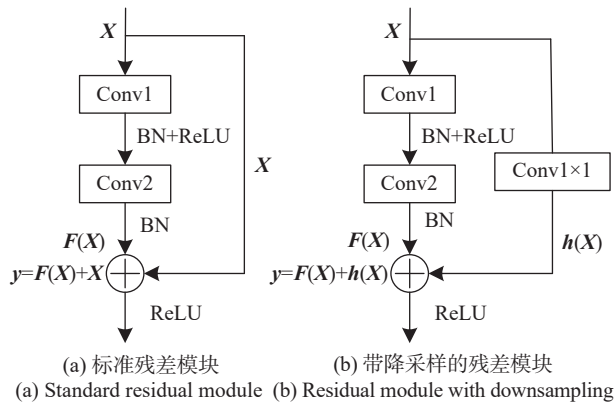


图 1 残差模块结构图

Fig. 1 Structural diagram of residual module

由图 1(a) 可见, 输入 X 分别经过第 1 个卷积层 Conv1、第 2 个卷积层 Conv2 以及批归一化 BN、激活函数 ReLU 得到残差映射 $F(X)$, 残差模块的期望输出 y 为 $F(X)$ 和残差模块输入 X 的和。当输入与残差映射输出的特征维度不同时, 残差模块结构如图 1(b) 所示, 需要在标准残差模块的基础上增加 1×1 的卷积运算 Conv1 $\times 1$ 得到输出 $h(X)$, 从而保证它们具有相同的特征维度。

1.3 SE 注意力模块

SE 注意力模块是从特征通道入手, 通过学习方式获取每个通道的重要程度, 并赋予不同的权重, 从

而突出重要特征, 抑制次要特征^[24]。该模块由压缩和激励两部分组成, 压缩操作采用全局平均池化函数, 在空间维度上对每一个特征通道进行压缩, 其计算公式如下:

$$\mathbf{Z} = \mathbf{F}_{sq}(\mathbf{X}_C) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \mathbf{X}_C(i, j) \quad (3)$$

式中, \mathbf{Z} 为压缩部分生成的权重; $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_C]$, \mathbf{X}_C 为第 C 个通道的特征图; $\mathbf{F}_{sq}(\cdot)$ 为对特征通道的压缩操作; H 和 W 分别表示特征图的长和宽。

激励操作用来获取每个特征通道的权重值, 计算公式如下:

$$\mathbf{S}_C = \mathbf{F}_{ex}(\mathbf{Z}, \mathbf{w}) = \sigma(\mathbf{w}_2 \text{ReLU}(\mathbf{w}_1 \mathbf{Z})) \quad (4)$$

式中, $\mathbf{w}_1 \in \mathbf{R}^{r \times C}$ 和 $\mathbf{w}_2 \in \mathbf{R}^{C \times r}$ 分别为第一层和第二层的全连接网络参数, 其中 r 为降维系数; $\mathbf{F}_{ex}(\cdot)$ 为对特征通道的激励操作; \mathbf{S}_C 为第 C 个通道的权重值; σ 为 Sigmoid 激活函数。

最后, 通过 $\mathbf{F}_{scale}(\cdot)$ 操作进行逐通道加权, 计算公式如下:

$$\overline{\mathbf{X}}_C = \mathbf{F}_{scale}(\mathbf{X}_C, \mathbf{S}_C) = \mathbf{X}_C \otimes \mathbf{S}_C \quad (5)$$

式中, $\overline{\mathbf{X}}_C$ 为加权后的特征输出, $\overline{\mathbf{X}} = [\overline{\mathbf{X}}_1, \overline{\mathbf{X}}_2, \dots, \overline{\mathbf{X}}_C]$; \otimes 为通道方向的乘法。

1.4 元迁移学习

元学习是解决小样本问题的有效工具, 其目标是使模型获取一种学会学习的能力。与传统机器学习不同, 元学习的关注点不再聚焦于特定任务, 而是探索不同任务之间的学习能力, 核心思想是从多个相关任务中学习数学模型, 并通过少量样本即可快速适应新任务^[25-26]。元学习的数据集分为元训练集 $D_{\text{meta-train}}$ 和元测试集 $D_{\text{meta-test}}$, 且两数据集之间没有交集。从由元训练集构成的任务分布 $p(T)$ 中采样一系列的任务 T , 其中每个任务 T_i 中包含 N 个类别, 每个类别中包含 K 个样本, 这被称为 N 类 K 样本问题。每个问题又被分为支持集和查询集, 分别记为 T_i^s 和 T_i^q 。目前使用最广泛的是 MAML 方法, 其基本思想是从多个相关任务上不断学习, 以获得更好的初始化参数, 使得 $D_{\text{meta-test}}$ 在此参数基础上使用少量样本通过有限步更新就能适应新任务, 其原理如图 2 所示。

迁移学习是利用源域数据对网络进行预训练以获取良好的泛化性能, 进而将模型迁移至目标域以提升新任务学习效率。如图 3 所示, 该方法通过微调预训练模型的部分或全部参数, 使其更好地适应目标任务。

元迁移学习是将元学习和迁移学习结合起来的图片分类算法^[27]。该算法分为预训练和元学习两个阶段, 预训练阶段是利用源域数据对提取特征的网

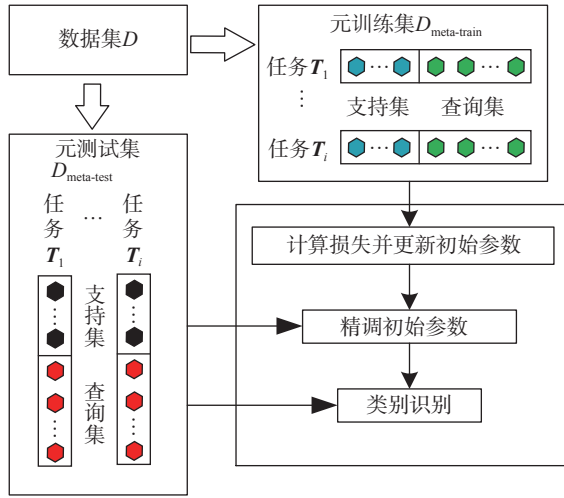


图 2 MAML 算法原理

Fig. 2 Principle of MAML algorithm

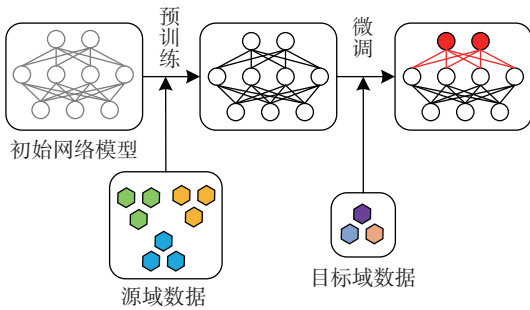


图 3 迁移学习原理

Fig. 3 Principle of transfer learning

络权重 θ 和分类器 θ 进行初始化, 并利用梯度下降法进行参数更新, 公式如下:

$$L_{D_{\text{meta-train}}}([\theta, \theta]) = \frac{1}{|D_{\text{meta-train}}|} \sum_{x_i, y_i \in D_{\text{meta-train}}} l(f_{[\theta, \theta]}(x_i), y_i) \quad (6)$$

$$[\theta, \theta] \leftarrow [\theta, \theta] - \alpha \nabla L_{D_{\text{meta-train}}}([\theta, \theta]) \quad (7)$$

式中, α 为预训练阶段学习率; $l(\cdot)$ 为交叉熵损失函数; $L_{D_{\text{meta-train}}}(\cdot)$ 为训练集的总体损失; ∇ 为计算梯度。

由于骨干网络的参数量过大, 若直接进行微调会增加过拟合的风险, 该算法在微调阶段设计了缩放和偏移(scale and shift, SS)操作, 减少了需要微调的参数量, 其原理如图 4 所示, 具体来说, 在进行参数迁移时, 将预训练好的网络参数 $\{w, b\}$ “冻结”(如

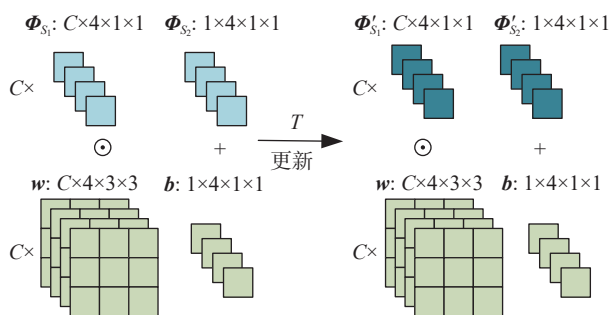


图 4 SS 操作的参数迁移

Fig. 4 Parameter transfer of the SS operation

图 2 下方部分所示), 并为“冻结”的参数添加可学习参数 Φ_{S_1} 和 Φ_{S_2} , 分别学习权重 w 的缩放以及偏置 b 的偏移以适应新的小样本学习任务。SS 操作公式如下:

$$SS(x; w, b; \Phi_{\{S_1, S_2\}}) = (w \odot \Phi_{S_1})x + (b + \Phi_{S_2}) \quad (8)$$

式中, \odot 表示逐元素相乘; x 为整个网络的输入。

元学习阶段是以元任务的方式展开, 将丢弃预训练阶段的分类器 θ , 通过内循环和外循环双层进行优化, 内循环是利用 $D_{\text{meta-train}}$ 中 T_i^s 通过梯度下降法优化对应任务 T_i 的分类器权重 θ_i , 参数更新如下:

$$\theta_i \leftarrow \theta - \beta \nabla_{\theta} L_{T_i^s}([\theta, \theta], \Phi_{\{S_1, S_2\}}) \quad (9)$$

式中, β 为内循环学习率; θ 为小样本学习分类器, 与预训练阶段的分类器不同; ∇_{θ} 为任务 T_i 在支持集上的损失 $L_{T_i^s}(\cdot)$ 的梯度。

外循环是根据 $D_{\text{meta-train}}$ 中的查询集 T_i^q 分别对参数 Φ_{S_1} 、 Φ_{S_2} 和 θ 进行梯度更新:

$$\Phi'_{S_1} \leftarrow \Phi_{S_1} - \gamma \nabla_{\Phi_{S_1}} \sum_{T_i \sim p(T)} L_{T_i^q}([\theta; \theta'], \Phi_{\{S_1, S_2\}}) \quad (10)$$

$$\Phi'_{S_2} \leftarrow \Phi_{S_2} - \gamma \nabla_{\Phi_{S_2}} \sum_{T_i \sim p(T)} L_{T_i^q}([\theta; \theta'], \Phi_{\{S_1, S_2\}}) \quad (11)$$

$$\theta \leftarrow \theta - \gamma \nabla_{\theta} \sum_{T_i \sim p(T)} L_{T_i^q}([\theta; \theta'], \Phi_{\{S_1, S_2\}}) \quad (12)$$

式中, Φ'_{S_1} 和 Φ'_{S_2} 分别为参数 Φ_{S_1} 和 Φ_{S_2} 的更新值; γ 为外循环学习率; $\nabla_{\Phi_{S_1}}$ 和 $\nabla_{\Phi_{S_2}}$ 分别为参数 Φ_{S_1} 和 Φ_{S_2} 的梯度; $L_{T_i^q}(\cdot)$ 为任务 T_i 在查询集上的损失。

以上描述为元训练过程, 对所有任务进行元训练后, 得到具有良好性能的元学习器, 以用于元测试。

2 本文所提方法

在实际工程中, 轴承的工况复杂多变且故障样本少, 一些小样本算法特征提取能力有限, 泛化能力较弱而难以应对。为此, 本研究引入 SE 注意力机制和残差网络, 借助元学习和迁移学习的优势, 设计了一种小样本下基于元迁移学习的变工况轴承故障诊断方法, 挖掘更深层次特征且提取关键的信息, 通过已知工况样本数据准确诊断出其余工况下轴承状态, 其流程图如图 5 所示。整个过程主要由信号预处理、模型预训练与元迁移训练以及变工况下的轴承故障诊断三个步骤组成。

2.1 信号预处理

使用加速度传感器采集多工况、多状态下的轴承振动信号, 并将采集到的一维时域信号无重叠地划分为若干等长的样本。由于时域信号很难发现特征, 且存在采样点密集、信号冗余等问题, 而时频图

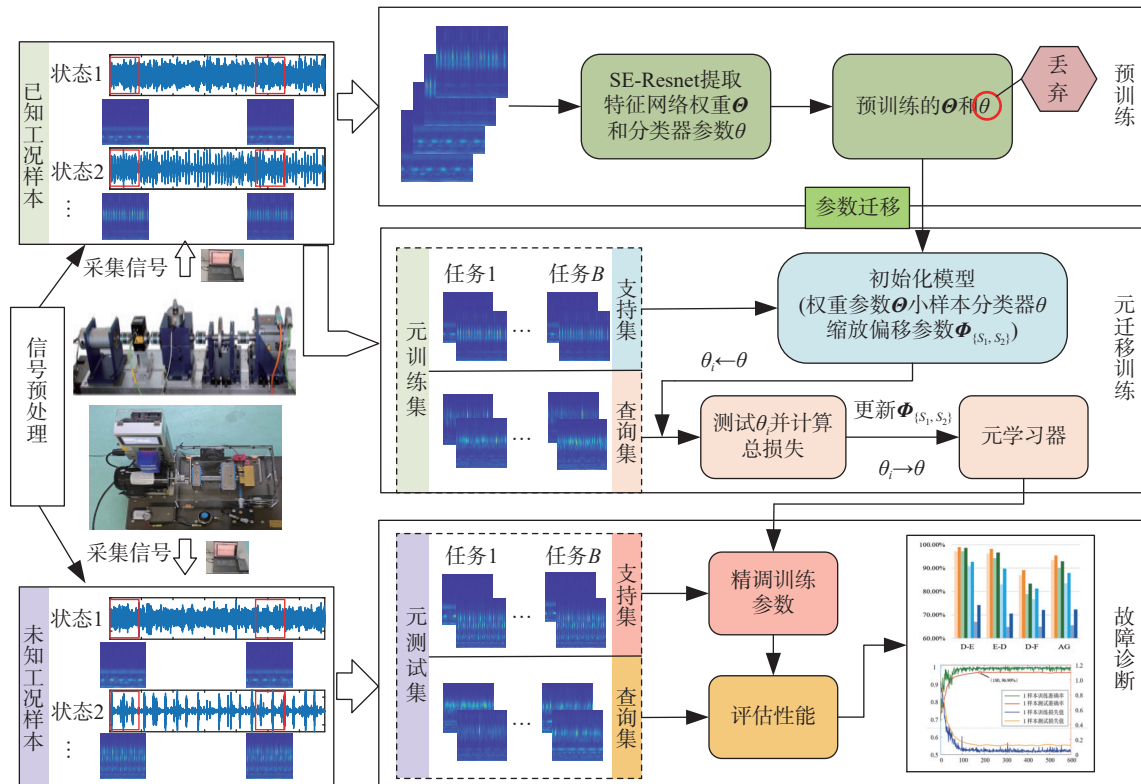


图 5 所提方法流程图

Fig. 5 Flowchart of the proposed method

可提供时频域上的特征, 包含的信息更突出、更直观, 有利于特征提取。故通过 CWT 将轴承一维振动信号转化为 RGB 三通道的时频图, 从而将轴承故障诊断转换为图像识别问题。按照上述步骤构造不同工况下的时频图, 并将已知工况下的样本作为训练集, 其余工况作为测试集, 完成数据集的构建。

2.2 模型预训练与元迁移训练

本文所构建的元迁移学习网络在 MAML 框架的基础上融合了迁移学习的思想, 而 MAML 通常被用在图像识别中, 其要求训练集和测试集之间没有交集。在图像分类中, 可根据图像识别内容属性获得大量类别, 然而在轴承的实际运行过程中, 轴承的故障类型有限, 若仅仅以故障状态为判别标准很难实现元任务的构建。考虑到轴承运行条件不断变化, 将工况条件也作为轴承故障类型的判别标准, 即将不同工况下的同种故障视为不同的故障类型。以已知工况的轴承样本为训练集, 新工况的故障样本为测试集, 进行变工况下的轴承故障诊断。另外, 为避免过拟合, MAML 的骨干网络通常为浅层的神经网络, 特征提取有限。残差网络可提取到大量的故障特征信息, 但其中也包含一些分布在某些通道上的无效特征信息, 这分散了网络的“注意力”, 从而影响了故障诊断的精度。故本文引入 SE 注意力和 ResNet 模块, 构建了一种 SE-ResNet 网络结构作为骨干网络, 其结构和模块如图 6 所示。一方面可通过

深度残差模块避免网络退化问题, 另一方面接入 SE 模块可保留有用信息, 抑制无用信息。

预训练阶段: 通过训练集的数据对骨干网络模型 SE-ResNet 进行训练, 利用式 (6) 和 (7) 不断更新优化特征提取权重 θ 和分类器参数 θ , 为后续的元迁移训练提供良好的初始化参数, 从而避免在利用少样本进行元训练时出现过拟合。

元迁移训练阶段: 该阶段丢弃预训练得到的分类器 θ , 保留特征提取权重 θ , 以任务驱动的方式进行训练。首先, 从训练集的所有故障类型中选择 N 类故障, 且每类故障有 K 个样本作为支持集, 剩余样本作为查询集, 从而生成一个元训练任务, 每批共有 B 个元任务组成。其次, 沿着不同元任务 (即已知工况下不同故障场景) 的聚合梯度下降方向来优化模型参数, 其计算公式如式 (12) 所示, 需要注意的是, 特征提取权重 θ 保持不变, 通过缩放和偏移参数 ϕ_{s_1} 和 ϕ_{s_2} 调整网络参数, 二者更新如式 (10) 和 (11) 所示; 每个元任务的参数自适应如式 (9) 所示; 最后, 通过多批先验任务的不断训练得到使用少样本微调就能达到高精度的元迁移网络。

2.3 变工况下轴承故障诊断

随机采集不同工况下的轴承状态数据并转换为二维时频图。与元训练任务的生成规则相同, 生成元测试任务 (新工况下不同故障场景), 由元测试的支持集在训练得到的元迁移网络基础上进行几步更

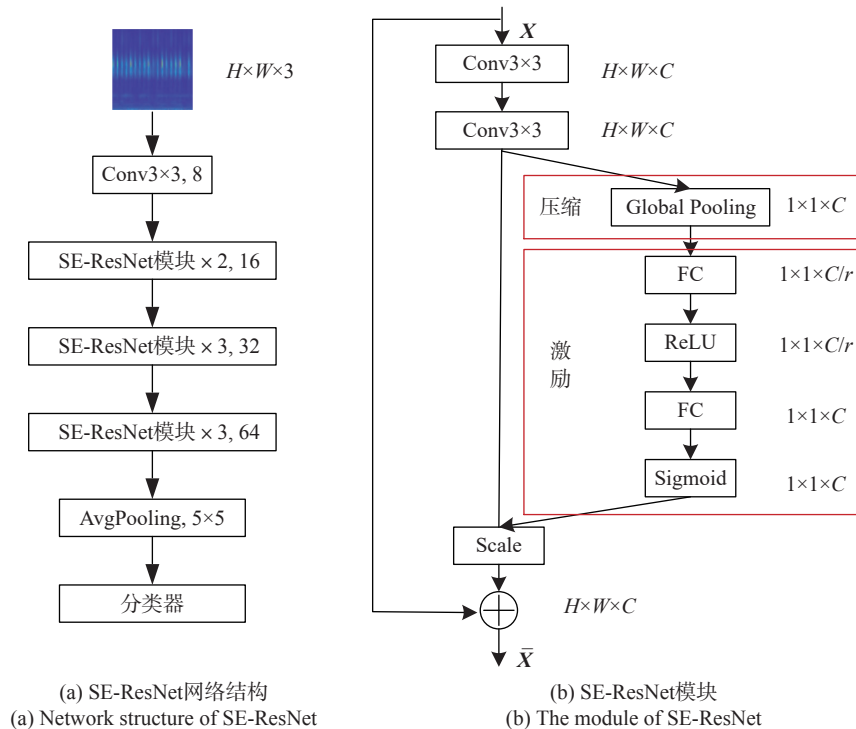


图 6 SE-ResNet 的网络结构和模块

Fig. 6 Network structure and module of SE-ResNet

新获得对应新任务的最优参数,由查询集执行故障诊断。

3 试验验证

为验证本文方法在小样本、变工况下轴承故障的诊断能力,本节选用实际工况下的轴承数据(美国凯斯西储大学(CWRU)和德国帕克伯恩大学(PU)轴承数据,以及实验室搭建的轴承故障模拟试验台采集数据)进行分析,并与其他小样本学习方法对比,验证所提方法的有效性和优越性。试验采用了搭载 AMD Ryzen 7 @3.20 GHz, 英伟达 RTX 3050 Ti 显卡的计算机系统。

3.1 CWRU 数据集

凯斯西储大学轴承试验台主要由驱动电机、扭矩传感器和功率测试机等组成。轴承型号为 6205-2RS JEM SKF, 采样频率为 12 kHz, 传感器布置在驱动端位置, 轴承缺陷均为电火花加工, 选取损伤程度分别为 0.1778、0.3556 和 0.5334 mm 不同量级的轴承内圈故障、外圈故障、滚动体故障和正常状态共 10 类轴承状态进行分析。

3.1.1 信号预处理

采集 10 类轴承状态在 1797、1772、1750 和 1730 r/min 转速下, 对应负载为 0、1、2 和 3 hp 的振动数据, 并按照工况条件将其分别标记为工况 Z、A、B 和 C, 详细设置如表 1 所示。相关研究表明, 同种故

障在不同工况下的振动数据具有一定的差异, 因此, 将 4 种工况下的同种故障类型视为不同类别。每类状态信号从第一个点开始取连续的 2048 个点作为一个样本, 进行不重叠采样, 样本采样过程如图 7 所示。

表 1 CWRU 数据集的工况设置

Tab. 1 Working conditions setting of CWRU dataset

电机转速/ ($r \cdot \min^{-1}$)	负载/hp	故障尺寸/ mm	故障类型	工况
1797	0	0.1778;	正常状态; 内圈故障; 外圈故障; 滚动体故障	Z
1772	1	0.3556;		A
1750	2	0.5334		B
1730	3			C

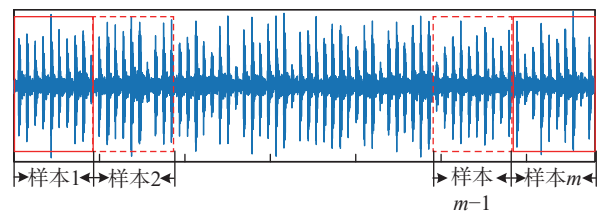


图 7 样本采样过程

Fig. 7 Sample sampling process

每类轴承状态选取 20 个样本, 每种工况共有 200 个样本。为了能够充分提取轴承振动信号的隐形特征, 以 $\text{cmor}1.5-2$ 为小波基函数, 对轴承原始振动信号进行 CWT 变换, 将其转换为大小为 84×84 的三通道时频图像。以工况 B, 损伤程度为 0.3556 mm 的内圈故障、外圈故障、滚动体故障以及正常状态为例, 其原始信号及时频图如图 8 所示。

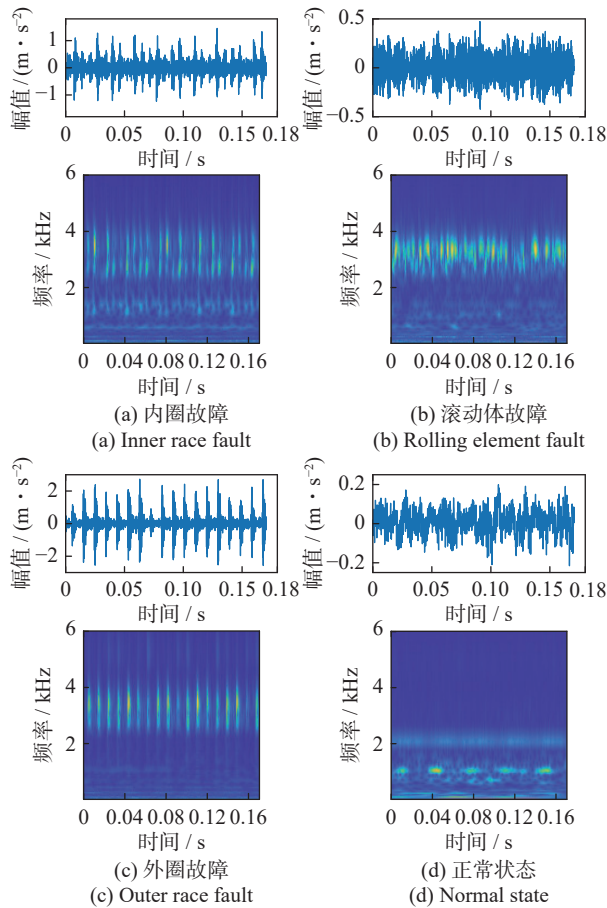


图 8 轴承原始信号及时频图

Fig. 8 Bearing original signal and time-frequency diagrams

3.1.2 模型训练与故障诊断

在工程实际中, 轴承常常运行在变转速、变负载条件下, 工况复杂, 难以获取每个工况下的轴承数据。模型能利用已知工况数据诊断出未知工况的轴承状态是保证滚动轴承安全运行的前提。随机选择 CWRU 数据集中一种工况进行训练, 其余工况进行测试, 以检验所提方法在小样本、变工况条件下的故障诊断性能, 即以工况 C 作为训练集, 以工况 Z、A 和 B 作为测试集, 分别记为 C-Z、C-A 和 C-B。

在元迁移训练和元测试阶段是以元任务的方式进行 (N 类 K 样本) 进行, 参考现有小样本研究方法的试验设置^[16-17], 将 N 设置为 5, K 设置为 1 或 5, 即每次随机从训练集中的 10 个类别中随机选取 5 类, 每个类别选择 1 或 5 个样本作为支持集, 15 个样本作为查询集, 如此生成一个元任务。每批共由 12 个元任务组成。模型的超参数设置如下: 预训练的批处理大小为 12, 预训练和元学习内循环学习率均为固定值, 设置为 0.001; 外循环初始学习率为 0.001 并采用等间隔调整学习率; 迭代次数为 600, 元迁移训练和元测试的更新步均设置为 10, 选用交叉熵损失函数反映模型预测标签与实际标签之间的差距, 整个训练采用 Adam 优化器。

在测试阶段, 元测试任务的生成与元训练任务

相同, 由支持集对已训练好的模型进行微调, 由查询集实施故障诊断。为了保证结果的稳定性, 反复生成元测试任务 50 次, 计算平均准确率且每种工况重复运行 3 次, 计算准确率的平均值。各工况组合下 1 样本和 5 样本的识别准确率如表 2 所示。可以看出, 5 样本下的准确率均高于 1 样本下的准确率, 其原因为 5 样本比 1 样本拥有更多的训练样本, 可为模型提供更全面的故障信息。还可以发现, 1 样本和 5 样本下所提方法在所有工况下的平均识别准确率 AG 均在 96% 以上, 这表明所提模型在小样本和变工况的条件下能够有效实现轴承故障诊断, 且具有良好的泛化性能。

表 2 所提方法在不同工况下的识别准确率

Tab. 2 Recognition accuracies of the proposed method under different working conditions

方法	识别准确率/%			
	C-Z	C-A	C-B	AG
所提方法-1样本	92.82	97.41	99.00	96.41
所提方法-5样本	95.96	98.40	99.48	97.95

3.1.3 CWRU 数据集方法对比

为了能证明所提方法在小样本和变工况条件下的优越性, 本文设置了 3 种变量, 即骨干网络模型、注意力机制和模型样本的输入类型, 对应 3 种小样本学习方法, 与本文方法进行对比分析。方法 1 为 MAML, 其骨干网络为只有 4 层卷积的浅层网络。方法 2 在所提方法的基础上取消了 SE 注意力机制, 记为 ResNet-MTL。方法 3 将轴承原始振动信号通过 STFT 转换为二维图像作为模型输入, 记为 STFT-MTL。其余参数设置均与本文方法相同, 各模型在不同工况组合下的识别准确率如图 9 所示。

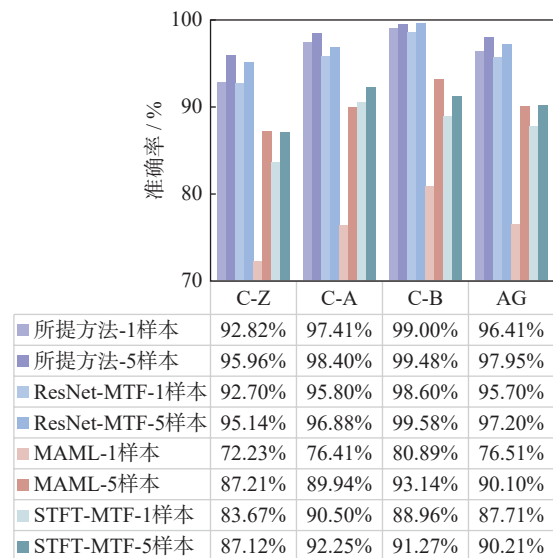


图 9 各模型在不同工况组合下的识别准确率

Fig. 9 Recognition accuracies of each model under different combinations of working conditions

由图 9 可知, 所提模型在相同样本下各工况组合的识别性能最佳。由于残差网络能够挖掘更深层次特征, 为模型提供更多有效的信息, 而 MAML 的骨干网络是浅层神经网络, 提取特征较少。另外, 元迁移学习通过预训练和元迁移训练两步学习策略, 具有快速适应能力, 使得所提方法和 ResNet-MTL 在所有工况下识别准确率都比 MAML 高。尤其在工况 C-A 下, 所提方法在 1 样本和 5 样本下分别比 MAML 高 21.00% 和 8.46%。与 ResNet-MTL 相比, 所提方法在 1 样本和 5 样本下 AG 均略高于该方法, 在工况 C-A 下, 所提方法 1 样本比其高 1.61%。而二者的唯一区别是 SE 注意力模块的引入, 这也表明 SE 模块能够使模型聚焦于更有效的特征通道, 抑制无效特征, 进而提高故障诊断的精度。与 STFT-MTL 相比, 所提方法在所有工况下的识别准确率均明显优于该方法, 在 1 样本和 5 样本下 AG 分别高 8.70% 和 7.74%。二者的唯一区别是模型输入类型, STFT-MTL 模型输入为由 STFT 生成的二维图片, 本文方法的输入为由 CWT 转换的时频图, 这也表明小波时频图能够为模型提供更具有表征性的特征信息, 有利于故障特征的学习。以上分析表明, 所提方法在小样本和变工况下具有更好的泛化能力和识别精度。

3.2 PU 数据集

为排除单一数据集对所提模型效果的影响, 本节选用德国帕德博恩大学不同工况下的轴承数据集进行测试验证, 其试验台如图 10 所示, 主要由电机、转矩仪、滚动轴承测试模块、飞轮和负载电机等组成。轴承型号为 6203 型滚动轴承, 通过电火花 (EDM)、钻孔 (Dg) 和电动雕刻 (EE) 的方法在轴承内圈和外圈施加不同程度 (level1 和 level2) 的损伤, 人工损伤的轴承故障类型如图 11 所示。选取 5 种外圈故障、3 种内圈故障以及正常状态共 9 种类型进行分析。

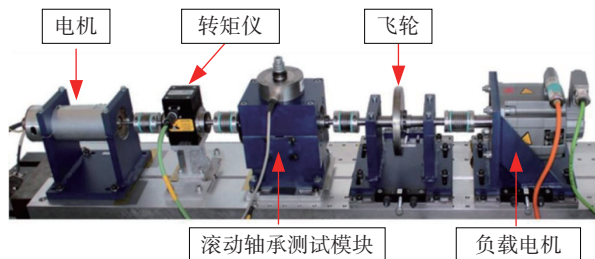


图 10 PU 数据集的轴承试验台

Fig. 10 Bearing test bench of PU dataset

3.2.1 信号预处理

以 64 kHz 的采样频率采集 9 类轴承状态在转速为 1500 r/min 时, 不同加载力矩和径向加载力下对应不同工况的振动信号, 并将其分别标记为工况 D、

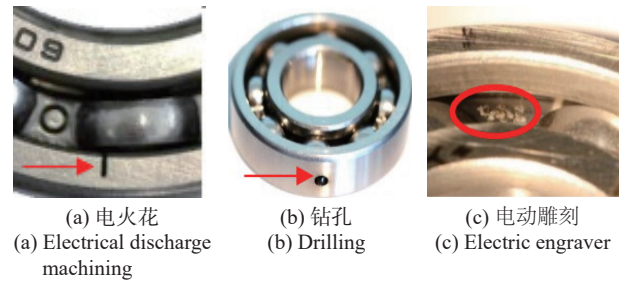


图 11 人工损伤的轴承故障类型

Fig. 11 Bearing fault types generated by artificial damage

E 和 F, 详细设置如表 3 所示。每类轴承故障的样本数为 20, 采样点数为 4096。如 3.1.1 节所述, 将每个故障样本转换为 84×84 的三通道时频图像。限于篇幅, 未将各工况下的时频图列出。

表 3 PU 数据集的工况设置

Tab. 3 Working conditions setting of PU dataset

转速/ ($r \cdot \min^{-1}$)	力矩/ ($N \cdot m^{-1}$)	径向力/ N	故障类型	工况
1500	0.1	1000	正常状态; OR-EDM(level 1); OR-EE(level 1); OR-EE(level 2);	D
1500	0.7	1000	OR-Dg(level 1); OR-Dg(level 2); IR-EDM(level 1);	E
1500	0.7	400	IR-EE (level 1); IR-EE (level 2)	F

3.2.2 模型训练与故障诊断

机械设备的运行工况复杂多变, 为了能够满足实际工况的需要, 且进一步验证所提模型在小样本和变工况下的有效性, 随机选择工况 D 作为训练集, 工况 E 和 F 为测试集进行试验验证, 分别记为 D-E 和 D-F。通过预训练和元迁移训练对模型进行训练, 元迁移训练和元测试均以元任务的形式进行, 元任务的生成方式和模型的参数设置如 3.1.2 节所述。为避免结果的随机性, 每种工况组合运行 3 次, 求准确率的平均值, 其识别结果如表 4 所示。可以发现, 由于 5 样本较 1 样本拥有更多的训练样本, 使得 5 样本的准确率均高于 1 样本。在工况 D-E 下 1 样本和 5 样本的准确率分别达到了 97.15% 和 98.92%, 且所有工况下的平均准确率 AG 在 92% 以上, 这表明在变工况下所提方法能通过有限的样本

表 4 所提方法在不同工况下的识别准确率

Tab. 4 Recognition accuracies of the proposed method under different working conditions

方法	识别准确率/%		
	D-E	D-F	AG
所提方法-1样本	97.15	86.87	92.01
所提方法-5样本	98.92	89.23	94.08

达到较高的故障识别准确率。

常规的小样本学习方法, 模型收敛之后随着迭代次数的增加, 很有可能出现过拟合现象。为了能够清晰展现出所提模型的收敛性能, 以工况 D-E 为例, 每训练 30 次进行 1 次测试, 其对应的准确率和损失值变化曲线如图 12 所示。从图 12(a)中可以看出, 在 1 样本下迭代次数达到 180 时, 测试准确率为 96.90%, 之后逐渐趋于收敛。从图 12(b)中可以看出, 模型在第 90 次迭代之后便逐渐趋于稳定状态, 二者对比可发现, 5 样本比 1 样本具有更快的收敛速度。另外, 二者在迭代过程中均未出现过拟合和欠拟合的现象, 其原因是元迁移学习的引入, 使得模型不但可以挖掘样本深层次特征, 还可以有效避免模型过拟合的发生。

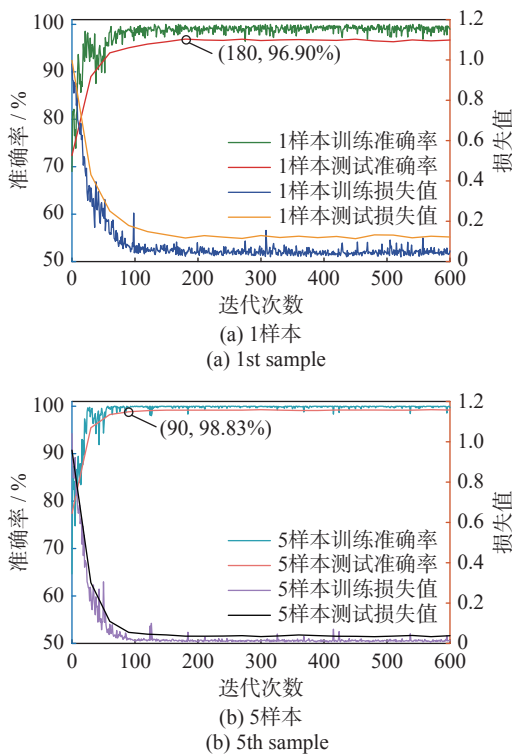


图 12 工况 D-E 下的准确率和损失值变化曲线

Fig. 12 Variation curves of accuracy and loss value under working conditions D-E

3.2.3 PU 数据集方法对比

为了验证所提模型在故障诊断中的优势。如 3.1.3 节所述, 将所提模型分别与 MAML、ResNet-MTL、STFT-MTL 进行对比分析。另外, 与一阶元学习算法 Reptile^[28] 对比分析, 其模型输入和参数设置均与本文方法相同。各模型在不同工况组合下的识别准确率如图 13 所示。可以看出, 所提方法拥有最高的识别准确率。与 STFT 相比, CWT 在时域频域都有良好的局域特性, 能够更全面地反映轴承振动特征信息, 使得以由 CWT 转换得到的时频图作为输入的模型的识别准确率均高于 STFT-MTL。残差网

络在避免网络退化的同时可以挖掘更深层次的复杂特征, 且元迁移学习能为模型提供良好初始参数, 具有快速学习的优势, 从而使得 ResNet-MTL 在 1 样本和 5 样本下所有变工况组合的平均识别准确率 AG 分别比 MAML 高 4.18% 和 3.98%。由于 SE 注意力机制的引入, 使得所提方法在 1 样本和 5 样本下 AG 分别比 ResNet-MTL 高 4.07% 和 3.10%。与一阶元学习算法 Reptile 相比, 所提方法在 1 样本和 5 样本下 AG 分别比其高 8.63% 和 3.71%。试验结果进一步验证了以 CWT 转换的时频图作为模型输入, 将 SE-ResNet 作为骨干网络, 基于元迁移学习构建的故障诊断模型能够在有限样本下更准确地识别变工况的轴承故障状态, 具有更强的泛化性能。

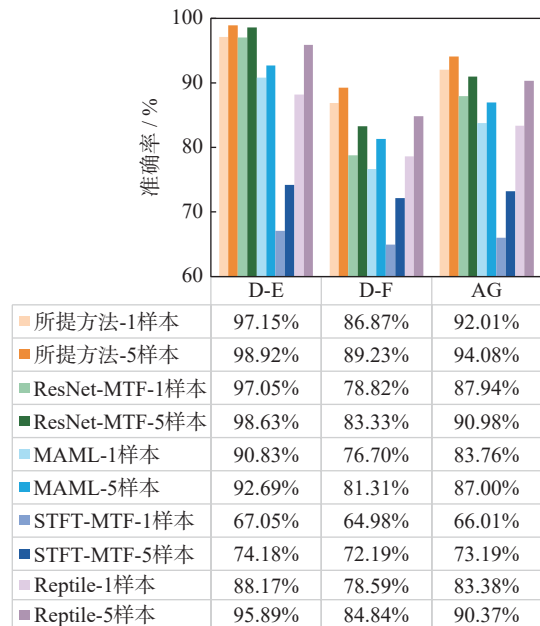


图 13 各模型在不同工况组合下的识别准确率

Fig. 13 Recognition accuracies of each model under different combinations of working conditions

3.3 实验室搭建的轴承故障模拟试验台

为了进一步验证所提方法的有效性及泛化能力, 通过实验室搭建的轴承故障模拟试验台采集轴承数据进行试验验证。试验平台如图 14 所示, 主要由交流电动机、转速控制器、转轴、支撑轴承、液压加载器和测试轴承等组成。测试轴承型号为 ER-16K, 采用激光刻蚀技术加工轴承缺陷, 数据集共模拟了 4 种轴承状态, 如图 15 所示, 分别为滚动体故障(故障直径为 1 mm)、内圈故障(故障宽度为 0.3 mm)、外圈故障(故障宽度为 0.3 mm)和正常状态, 其中滚动体为圆形故障, 内、外圈为矩形故障。

3.3.1 信号预处理

型号为 INV9821-140921 的 ICP 型加速度传感器安装在测试轴承座的正上方, 采集仪型号为 INV3062-

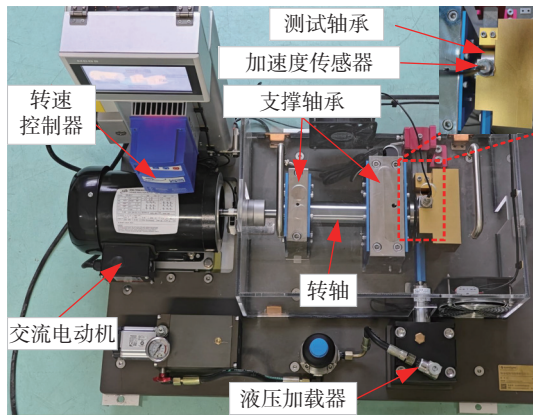


图 14 轴承故障模拟试验台

Fig. 14 Bearing fault simulation test bench



图 15 ER-16K 滚动轴承状态类型

Fig. 15 State types of ER-16K rolling bearing

C2, 以 12 kHz 的采样频率采集 10 s 四种轴承状态分别在 540 和 660 r/min 两种恒定转速, 540~660 和 660~540 r/min 两种变转速, 0、100、200 kg 三种负载组合下, 共 8 种工况下的振动数据并将其分别标记为工况 H、I、J、K、L、M、N 和 O, 详细设置如表 5 所示。每种轴承状态采用不重叠采样的方式进行采样, 采样点数为 2048, 样本数为 20。将轴承所有样本

表 5 实验室轴承数据集的工况设置

Tab. 5 Working condition setting of laboratory bearing dataset

电机转速/(r·min ⁻¹)	负载/kg	故障类型	工况
540	0		H
	100		I
	200	正常状态;	J
660	0	内圈故障;	K
	100	外圈故障;	L
	200	滚动体故障	M
540~600	100		N
660~540			O

利用 CWT 转换成大小为 84×84 的三通道时频图像。以工况 I 为例, 4 种轴承状态下的时频图如图 16 所示。

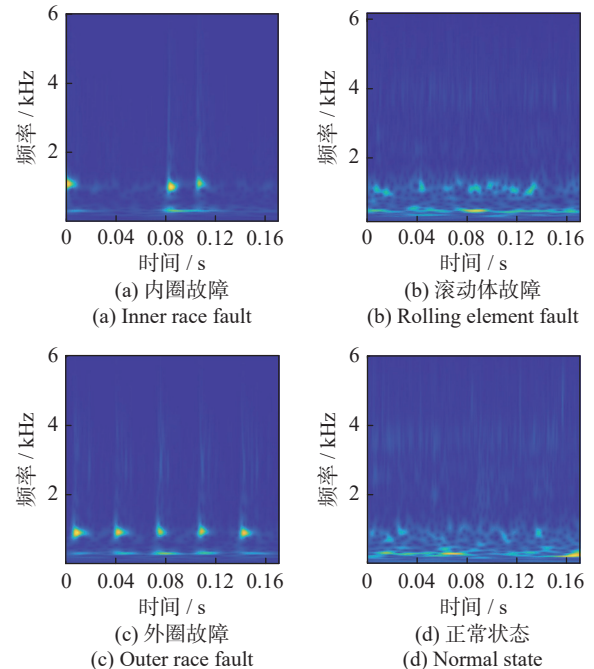


图 16 工况 I 下 4 种轴承状态的时频图

Fig. 16 Time-frequency diagrams of four bearing states under working condition I

3.3.2 模型训练与故障诊断

通过预训练和元迁移训练两步学习策略对模型进行训练, 元迁移训练和元测试仍按 N 类 K 样本的方式展开, 由于每种工况下的轴承状态类型只有 4 种, 故将 N 设置为 4, K 设置为 1 或 5。其余参数设置如下: 预训练学习率为 0.0004, 元学习内循环学习率为 0.001, 外循环初始学习率为 0.0001, 且并采用等间隔调整学习率。预训练和元任务的生成方式均与 CWRU 数据集相同, 如 3.1.1 节所述。一个优秀的模型能够利用有限的故障样本适应不同工况条件的变化, 为了能够进一步验证所提方法利用小样本实现复杂工况下轴承故障诊断的性能, 随机选用工况 I 为训练集, 其余工况为测试集进行测试。为展现工况 I 与其余工况下的数据分布差异, 参考文献 [29] 的方法, 采用以下步骤分析: 首先, 对各工况下轴承不同状态的振动信号进行快速傅里叶变换, 获取对应状态的频谱图; 其次, 对得到的频谱图利用高斯核密度估计求取对应状态下概率密度函数; 最后, 计算工况 I 与其余工况在同种轴承状态之间的 Kullback-Leibler (K-L) 散度值以量化表征工况之间的数据分布差异, 其结果如表 6 所示。可以发现, 工况 I 与工况 H、K 和 L 的 K-L 散度平均值较低, 表明工况 I 与它们之间的数据分布差异较小。工况 I 与工况 J、M、N 和 O 的 K-L 散度平均值较高, 表明工况 I 与它们之间的数据分布差异较大。

表 6 工况 I 与其余工况的数据分布差异

Tab. 6 Data distribution difference between working condition I and other working conditions

工况	外圈故障	内圈故障	滚动体故障	正常状态	平均值
H	0.1757	0.0883	0.1893	0.2571	0.1776
J	0.4221	0.2337	0.1624	0.2894	0.2769
K	0.0314	0.1234	0.1768	0.0358	0.0919
L	0.1516	0.0336	0.0243	0.2312	0.1102
M	0.4156	0.2093	0.1675	0.3874	0.2950
N	0.6273	0.4788	0.2746	0.3220	0.4257
O	0.6221	0.4821	0.3048	0.3063	0.4288

所提方法在不同工况组合下的识别准确率如图 17 所示。可以看出, 在工况 I-L 下, 1 样本和 5 样本的识别准确率达到 98% 以上。由于 5 样本可为模型提供更全面的样本信息, 使得 5 样本的识别准确率普遍高于 1 样本, 工况 I-N 下提升最高, 为 8.21%。在所有变工况下所提方法 1 样本和 5 样本的识别准确率均在 83% 以上。试验结果表明, 所提方法具有良好的泛化能力, 在有限的样本数量下, 能有效地判断新工况下轴承的运行状态。

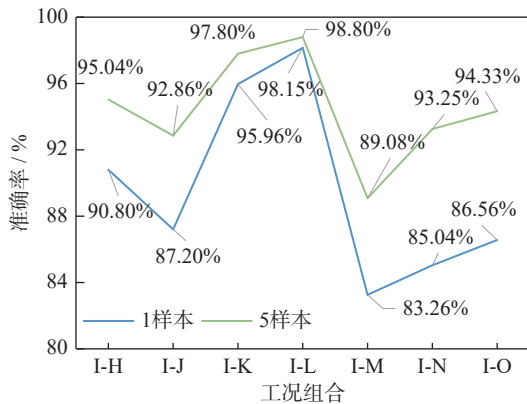


图 17 所提方法在不同工况组合下的识别准确率

Fig. 17 Recognition accuracies of the proposed method under different combinations of working conditions

为了能够直观地分析所提模型在时频图像中的关注特征, 以热力图的形式对网络提取的故障特征进行可视化展示。Grad-CAM 的核心机制是通过计算最后一个卷积层的梯度来确定各通道的影响权重, 并将其与特征图结合, 生成热力图, 并将该热力图与原图叠加来反映模型分类决策中每个区域的重要程度。本文采用 Grad-CAM 技术, 随机选择工况 K 下 1 个正常状态和 1 个外圈故障的样本进行热力图分析, 其结果如图 18 所示。在热力图中颜色越深的区域, 表示在故障诊断时, 模型对该区域的关注程度越高。可以发现, 时频图中频率幅值较高的区域, 均被所提模型重点关注, 这也表明所提模型能够捕获故障特征, 从而实现轴承故障状态的准确识别。

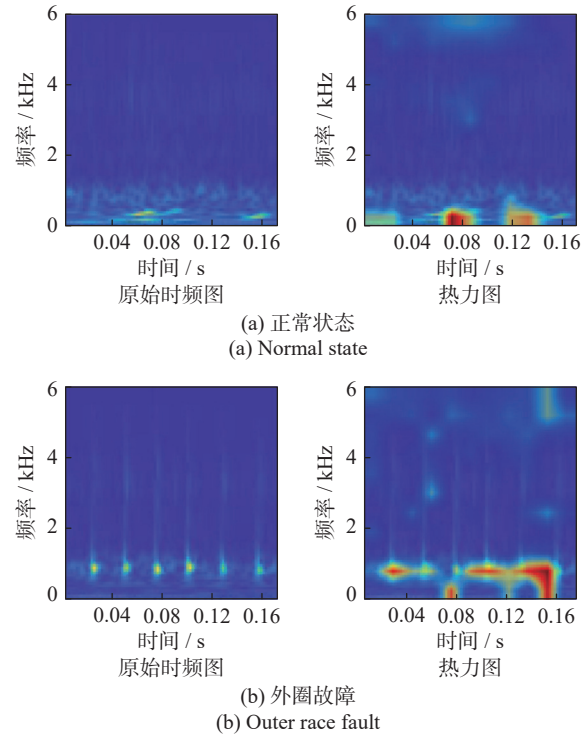


图 18 工况 K 下不同轴承状态的时频图和热力图

Fig. 18 Time-frequency diagrams and thermodynamic diagrams of different bearing states under working condition K

3.3.3 故障模拟试验方法对比

为了能够进一步验证所提方法在小样本、变工况下的优越性, 如 3.2.3 节所述, 分别采用 MAML、ResNet-MTL、STFT-MTL、Reptile 与本文方法进行对比分析。各模型在不同工况组合下的识别准确率如图 19 所示。可以看出, 所提方法在所有工况的同样样本条件下具有最佳故障诊断性能。由于 MAML 的骨干网络为浅层神经网络, 当工况差异较大时(如工况 I-N), 其无法为模型提供更多有效的特征信息。另外, 元迁移学习的预训练可为模型提供良好的初始化网络参数, 从而使得所提方法在 1 样本下的 AG 比 MAML 高 27.88%, ResNet-MTL 比 MAML 高 21.23%。就 AG 而言, 在 1 样本和 5 样本下所提方法分别比 ResNet-MTL 高 6.65% 和 7.05%。而二者唯一的不同在于所提方法引入 SE 注意力模块, 这也表明了 SE 注意力模块引入的必要性与有效性。在 1 样本和 5 样本下, 所提方法的 AG 分别比 STFT-MTL 高 22.32% 和 20.41%, 它们的不同在于模型输入类型。由 CWT 转换的时频图能够准确反映轴承运行状态, 可为模型提供更多有用信息。与一阶元学习算法 Reptile 相比, 所提方法 1 样本和 5 样本下的 AG 分别比其高 7.34% 和 2.60%。以上试验和分析结果进一步表明, 所提方法可利用有限样本数据实现轴承故障的准确识别, 且具有更强的泛化能力。

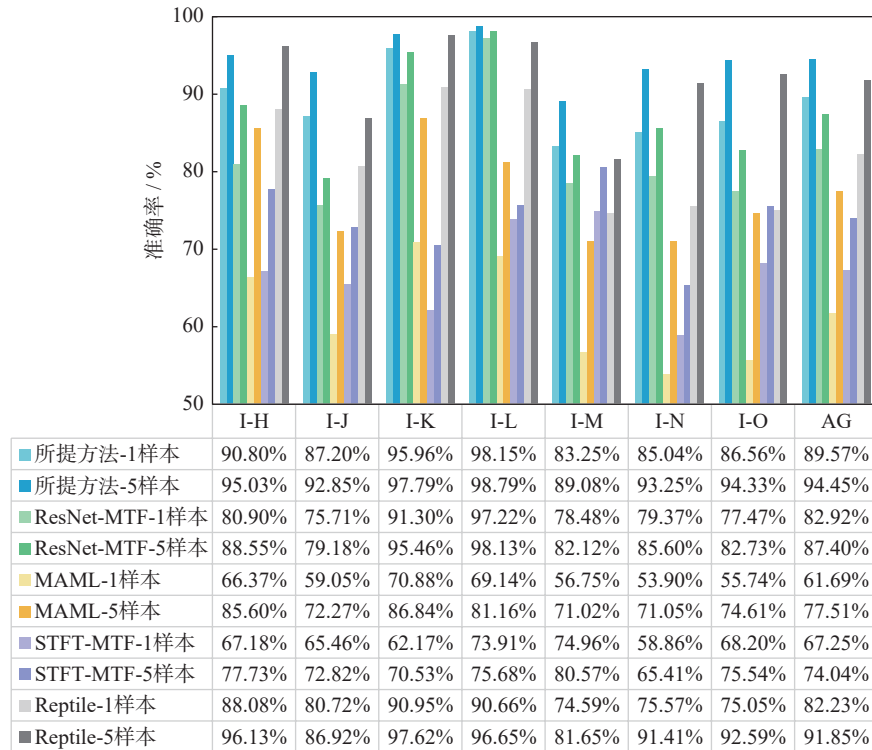


图 19 各模型在不同工况组合下的识别准确率

Fig. 19 Recognition accuracies of each model under different combinations of working conditions

4 结 论

针对滚动轴承在小样本和变工况条件下故障诊断精度不佳且泛化能力弱的问题,提出了小样本下 SE-ResNet 与元迁移学习的变工况轴承故障诊断方法。运用凯斯西储大学和帕德博恩大学两个基准轴承数据集以及实验室搭建的轴承故障模拟试验台采集数据进行试验验证。得出结论如下:

(1)通过 CWT 将一维轴承振动信号转换为二维时频图像,能更准确且充分地表征轴承故障特征信息,更有利于模型的特征提取与学习。

(2)将 SE-ResNet 作为骨干网络,在加深网络深度以避免梯度消失问题的同时能突出重要特征,消除信息冗余,能从有限样本中提取更具判别力的故障特征信息。

(3)结合元学习和迁移学习的各自优势,利用已知工况数据对网络模型进行预训练和元迁移训练,能够得到具有较强泛化能力的元迁移网络模型,通过少量样本即可实现变工况下端到端的轴承故障诊断,提高了模型的泛化能力。

参考文献:

[1] CHEN G Y, TANG G, ZHU Z X. VKCNN: an interpretable variational kernel convolutional neural network for rolling bearing fault diagnosis[J]. *Advanced Engineering Informatics*, 2024, 62: 102705.

[2] 王照伟, 刘传帅, 赵文祥, 等. 多尺度多任务注意力卷积

神经网络滚动轴承故障诊断方法[J]. *电机与控制学报*, 2024, 28(7): 65-76.

WANG Zhaowei, LIU Chuanshuai, ZHAO Wenxiang, et al. Rolling bearing fault diagnosis with multi-scale multi-task attention convolutional neural network[J]. *Electric Machines and Control*, 2024, 28(7): 65-76.

[3] KANG S Q, TANG X, WANG Y J, et al. Cross-domain fault diagnosis method for rolling bearings based on contrastive universal domain adaptation[J]. *ISA Transactions*, 2024, 146: 195-207.

[4] NIE G C, ZHANG Z W, JIAO Z H, et al. A novel intelligent bearing fault diagnosis method based on image enhancement and improved convolutional neural network[J]. *Measurement*, 2025, 242: 116148.

[5] 赵敬娇, 赵志宏, 杨绍普. 基于残差连接和 1D-CNN 的滚动轴承故障诊断研究[J]. *振动与冲击*, 2021, 40(10): 1-6.

ZHAO Jingjiao, ZHAO Zhihong, YANG Shaopu. Rolling bearing fault diagnosis based on residual connection and 1D-CNN[J]. *Journal of Vibration and Shock*, 2021, 40(10): 1-6.

[6] 于洋, 马军, 王晓东, 等. 基于 GST 与改进 CNN 的滚动轴承智能故障诊断[J]. *铁道科学与工程学报*, 2022, 19(7): 2050-2060.

YU Yang, MA Jun, WANG Xiaodong, et al. Intelligent fault diagnosis of rolling bearings based on GST and improved CNN[J]. *Journal of Railway Science and Engineering*, 2022, 19(7): 2050-2060.

[7] 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 Shouqiang, LIU Zhe, WANG Yujing, 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.
- [8] 赵一楠, 刘昌峰, 孟佳东, 等. 自适应窗口旋转优化短时傅里叶变换的变转速滚动轴承故障诊断[J]. 振动工程学报, 2024, 37(6): 1064-1076.
- ZHAO Yanan, YAN Changfeng, MENG Jiadong, et al. Fault diagnosis of rolling bearings under variable speed conditions based on adaptive window rotation optimization short-time Fourier transform[J]. Journal of Vibration Engineering, 2024, 37(6): 1064-1076.
- [9] XU Y, LI Z X, WANG S Q, et al. A hybrid deep-learning model for fault diagnosis of rolling bearings[J]. *Measurement*, 2021, 169: 108502.
- [10] CHENG Y W, LIN M X, WU J, et al. Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network[J]. *Knowledge-Based Systems*, 2021, 216: 106796.
- [11] 安文杰, 陈长征, 田森, 等. 基于迁移学习的风电机组轴承故障诊断研究[J]. 太阳能学报, 2023, 44(6): 367-373.
- AN Wenjie, CHEN Changzheng, TIAN Miao, et al. Research on bearing fault diagnosis of wind turbines based on transfer learning[J]. Acta Energetica Solaris Sinica, 2023, 44(6): 367-373.
- [12] 侯东晓, 穆金涛, 方成, 等. 基于 GADF 与引入迁移学习的 ResNet34 对变速轴承的故障诊断[J]. 东北大学学报(自然科学版), 2022, 43(3): 383-389.
- HOU Dongxiao, MU Jintao, FANG Cheng, et al. Fault diagnosis of variable speed bearings based on GADF and ResNet34 introduced transfer learning[J]. Journal of Northeastern University (Natural Science), 2022, 43(3): 383-389.
- [13] WANG X, SHEN C Q, XIA M, et al. Multi-scale deep intra-class transfer learning for bearing fault diagnosis[J]. Reliability Engineering & System Safety, 2020, 202: 107050.
- [14] YANG T Y, TANG T, WANG J W, et al. A novel cross-domain fault diagnosis method based on model agnostic meta-learning[J]. *Measurement*, 2022, 199: 111564.
- [15] LIN J, SHAO H D, MIN Z S, et al. Cross-domain fault diagnosis of bearing using improved semi-supervised meta-learning towards interference of out-of-distribution samples[J]. *Knowledge-Based Systems*, 2022, 252: 109493.
- [16] FINN C, ABBEEL P, LEVINE S, et al. Model-agnostic meta-learning for fast adaptation of deep networks[C]// Proceedings of the 34th International Conference on Machine Learning. ACM, 2017: 1126-1135.
- [17] CHEN J J, HU W H, CAO D, et al. A meta-learning method for electric machine bearing fault diagnosis under varying working conditions with limited data[J]. IEEE Transactions on Industrial Informatics, 2023, 19(3): 2552-2564.
- [18] SU H, XIANG L, HU A J, et al. A novel method based on meta-learning for bearing fault diagnosis with small sample learning under different working conditions[J]. *Mechanical Systems and Signal Processing*, 2022, 169: 108765.
- [19] HU Y D, LIU R N, LI X L, et al. Task-sequencing meta learning for intelligent few-shot fault diagnosis with limited data[J]. IEEE Transactions on Industrial Informatics, 2022, 18(6): 3894-3904.
- [20] SUN Q R, LIU Y Y, CHUA T S, et al. Meta-transfer learning for few-shot learning[C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019: 403-412.
- [21] YAN R Q, SHANG Z G, XU H, et al. Wavelet transform for rotary machine fault diagnosis: 10 years revisited[J]. *Mechanical Systems and Signal Processing*, 2023, 200: 110545.
- [22] 魏焕卫, 宋志鑫, 雷树立, 等. 基于残差网络的风电机组基础健康监测数据修复研究[J]. 太阳能学报, 2024, 45(4): 143-150.
- WEI Huanwei, SONG Zhixin, LEI Shuli, et al. Research on health monitoring data restoration of wind turbine foundation based on residual network[J]. Acta Energetica Solaris Sinica, 2024, 45(4): 143-150.
- [23] NI Q, JI J C, HALKON B, et al. Physics-Informed Residual Network (PIResNet) for rolling element bearing fault diagnostics[J]. *Mechanical Systems and Signal Processing*, 2023, 200: 110544.
- [24] FENG Y, CHEN J L, ZHANG T C, et al. Semi-supervised meta-learning networks with squeeze-and-excitation attention for few-shot fault diagnosis[J]. *ISA Transactions*, 2022, 120: 383-401.
- [25] 刘鑫磊, 冯林, 廖凌湘, 等. 基于元学习的图卷积网络少样本学习模型[J]. 电子学报, 2024, 52(3): 885-897.
- LIU Xinlei, FENG Lin, LIAO Lingxiang, et al. Few-shot learning on graph convolutional network based on meta learning[J]. Acta Electronica Sinica, 2024, 52(3): 885-897.
- [26] FENG Y, CHEN J L, XIE J S, et al. Meta-learning as a promising approach for few-shot cross-domain fault diagnosis: algorithms, applications, and prospects[J]. *Knowledge-Based Systems*, 2022, 235: 107646.
- [27] FAN L L, CHEN B Y, ZENG X J, et al. Knowledge-enhanced meta-transfer learning for few-shot ECG signal classification[J]. Expert Systems with Applications, 2025, 263: 125764.
- [28] LUO J J, SHAO H D, LIN J, et al. Meta-learning with elastic prototypical network for fault transfer diagnosis of bearings under unstable speeds[J]. Reliability Engineering & System Safety, 2024, 245: 110001.
- [29] 彭珍瑞, 刘臻. 基于故障可诊断性的齿轮箱传感器优化布置[J]. 振动与冲击, 2021, 40(4): 155-163.
- PENG Zhenrui, LIU Zhen. Optimal sensor placement of a gear box based on fault diagnosability[J]. Journal of Vibration and Shock, 2021, 40(4): 155-163.

第一作者:刘 臻(1994—),男,博士研究生。

E-mail: 2945341335@qq.com

通信作者:彭珍瑞(1972—),男,博士,教授。

E-mail: pzrui@163.com