

DOI: 10.13382/j.jemi.B2306246

多小波系数增强动态聚合联邦深度网络的 多工况故障诊断*

张焱^{1,2} 何姝钡^{1,2} 韩延^{1,2} 黄庆卿^{1,2}

(1. 重庆邮电大学工业物联网与网络化控制教育部重点实验室 重庆 400065;
2. 重庆邮电大学工业互联网研究院 重庆 401120)

摘要:针对分布式场景下单节点样本有限、多节点间工况分布不平衡等导致的深度学习故障诊断精度低的问题,提出一种多小波系数增强动态聚合联邦深度网络用于分布式小样本下的多工况机械故障诊断。提出多小波系数增强动态聚合联邦深度网络的诊断框架,单终端节点从本地样本中提取小波系数特征,提出多小波系数深度网络融合的特征增强方法,局部模型从多样性小波系数集合中提取更具判别性故障特征;聚合节点通过对多终端节点局部模型的聚合以构建全局联邦深度网络模型,并用于多工况故障诊断;为降低多节点间数据非独立同分布的影响,提出平衡模型贡献度的联邦动态加权聚合算法。轴承振动数据分析结果表明,所提方法能在分布式小样本条件下实现高精度的多工况故障诊断。

关键词: 故障诊断;小样本;多工况;联邦学习;特征增强

中图分类号: TH165.3;TN911.2 **文献标识码:** A **国家标准学科分类代码:** 510.40

Multi-wavelet coefficients enhanced dynamic aggregation federal deep network for fault diagnosis under multiple conditions

Zhang Yan^{1,2} He Shubei^{1,2} Han Yan^{1,2} Huang Qingqing^{1,2}

(1. Key Laboratory of Industrial Internet of Things & Networked Control, Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2. Institute of Industrial Internet, Chongqing University of Posts and Telecommunications, Chongqing 401120, China)

Abstract: To address the low accuracy of deep learning based fault diagnosis under distributed scenarios that caused by limited sample of single node and unbalanced distribution of working conditions of multiple nodes, et al, a multi-wavelet coefficients enhanced dynamic aggregation federal deep network (MWCE-FedDWA) is proposed for fault diagnosis under multiple conditions with distributed small samples. A framework for fault diagnosis using MWCE-FedDWA is proposed, wavelet coefficient features are extracted by each terminal node from its local samples, a method based on multi-wavelet coefficient fusion in deep network is proposed for feature enhancement, each local model utilizes a set of diversified wavelet coefficients to extract more discriminative fault features. A global federal deep network model is constructed in aggregation node by aggregating the local models from multiple terminal nodes, and then adopted for fault diagnosis under multiple conditions. To reduce the influence of non-independent and identically distributed data among multiple nodes, a federated dynamic weighted aggregation algorithm is proposed to balance the contribution of local models. The results on bearing vibration data show that the proposed method can achieve high-precision diagnosis under multiple conditions with distributed small samples.

Keywords: fault diagnosis; small sample; multiple conditions; federated learning; feature enhancement

0 引言

针对重大旋转机械的状态监测与故障诊断研究对于保障装备安全可靠运行具有重要意义^[1]。机械故障诊断本质上是模式识别问题,其中状态特征提取与故障辨识是其关键性步骤^[2]。传统机械故障诊断技术尤其是基于信号分析处理的故障诊断方法(如频谱分析^[3]、包络解调分析^[4]、小波变换^[5]和变分模态分解^[6]等),仅需对小样本(甚至单一样本)数据中的故障相关特征直接提取和辨识即可完成故障诊断。该类方法虽然极大地降低了对样本数的需求,但存在故障特征提取依赖于工程经验、以及固定模式特征提取难以适应变工况应用等局限。近年来,深度学习在许多分类识别任务中取得了显著成果,基于深度学习的故障诊断也逐步成为前沿性研究方向。然而,虽然深度学习具有强大的非线性表示和数据信息挖掘等方面能力,但基于深度学习的故障诊断其性能严重依赖于大数量、高质量(如类别平衡、恒定工况等)样本。

在工程实际中,大数量且高质量的数据集获取往往代价昂贵甚至难以实现。首先,单个数据拥有方拥有的数据样本量往往十分有限,分布式单终端节点的小样本量难以匹配深度学习大样本量的需求;其次,实际旋转机械往往在多变工况(如变转速、变负载等)条件下运行^[7],构建覆盖所有工况条件的数据集几乎不可能,常规深度学习故障诊断模型在多工况下泛化能力较差,同时,变工况运行也进一步加剧单终端节点特定工况下大样本集的构建难度;再次,近年来数据作为企业重要资产的观念逐步得到普遍认同,数据隐私保护需求提升进一步增大了“集中式”大样本集构建难度^[8]。上述因素制约了以大数据为基础的深度学习故障诊断实际应用。

联邦学习(federated learning, FL)^[9]作为一种新型的分布式机器学习技术,能够在数据隐私保护前提下利用多个数据拥有方的数据集以协作方式训练模型,为分布式小样本条件下基于深度学习的多工况故障诊断提供了新思路。有学者将联邦学习引入了故障诊断领域,Zhang等^[10]提出一种基于联邦学习和卷积神经网络的故障诊断方法,为故障诊断领域提供了一种有前景的协同训练方法;Zhou等^[11]设计了一种用于多工况轴承故障诊断的模块化联邦学习网络,利用动态路由技术改善了联邦学习的权值分配方案,但针对节点内部故障特征信息缺乏未做改进;Zhang等^[12]提出了基于动态验证和自我监督的联邦学习机械故障诊断在局部模型训练步骤中,人为地创造假样本,从有限的局部训练数据中学习结构知识,并使用验证集对局部模型进行评估,但这需要额外数据进行辅助诊断。同时,信号处理方法能够清晰表征信号

的物理属性,深度学习方法能够提取高度抽象化且具有极强表征能力的深度表达特征。有学者将信号处理方法同深度学习相结合,如魏秀业等^[13]提出一种基于特征融合与深度残差网络的行星齿轮箱故障诊断方法,利用模态分解方法并筛选出有效特征进行融合从而提高故障诊断精度;毛向向等^[14]将振动信号进行时频变换转换为时频图构成神经网络输入样本,提供了一种故障诊断新途径。然而在数据隐私保护前提下联邦学习和信号处理相结合的相关研究仍是空白的,同时,现有的联邦故障诊断方法虽考虑了联邦学习框架下多工况引起的数据分布问题,并未考虑到联邦深度学习模型训练时容易出现的小样本问题,而小样本问题在单节点故障诊断中是普遍存在的。值得庆幸的是,联邦学习中的局部模型训练原理与深度学习中单节点模型训练相似,因此利用信号处理方法来解决联邦局部模型小样本问题是可行的。

为解决上述问题,提出了一种多小波系数增强动态聚合联邦深度网络(multi-wavelet coefficients enhanced dynamic aggregation federal deep network, MWCE-FedDWA)并用于多工况故障诊断,该方法以联邦协作方式利用多个数据拥有方的数据集完成深度模型训练,无需多节点间源数据共享,契合大数据时代隐私保护需求下的故障诊断应用。本文主要工作包括:首先,为充分发挥传统基于信号分析处理的故障诊断方法对样本需求低、以及深度学习强大非线性表示和数据信息挖掘能力等优势,提出多小波系数深度网络融合的特征增强方法,终端节点联合多小波基函数对原信号进行分解以增强信号特征多样性,局部模型对多小波系数集合进行深度融合以自动提取更具判别性故障特征。其次,利用联邦学习技术联合拥有不同工况下样本的多节点进行联邦深度网络模型训练,以得到能匹配更多工况的诊断模型,针对因局部数据分布差异造成的模型诊断精度低的问题,提出一种新的平衡模型贡献度的联邦动态加权聚合算法,并用于聚合构建全局联邦深度网络模型。最后,利用齿轮箱轴承振动信号对所提方法进行了验证。

1 分布式场景下联邦学习故障诊断基础

联邦学习是一种新颖的分布式机器学习模型,可以当数据处于本地的情况下,进行数据共享,无需传递数据,对机械数据具有良好地检测能力。基于联邦学习独特的分布式学习优势,将联邦学习应用于机械故障诊断以解决机械装备时代故障数据不易共享及单个设备故障数据单一等问题。

基于联邦学习的故障诊断基础框架如图1所示,其包含两种类型的节点:聚合节点和终端节点。终端节点表示数据的所有者和局部模型的生产者,作为联邦训练

任务的发起者和参与者,负责训练自己的本地模型。聚合节点可以是云端服务器,作为受信任的第三方,负责模型的加密、解密和聚合以得到全局模型。

假设终端节点之间存在“数据孤岛”问题,即终端节点之间无法相互传输数据,终端节点和聚合节点分别使用本地化的训练算法和聚合算法,典型的模型训练过程如下:

1) 全局模型初始化:服务器确定训练任务以及相应的数据需求,并指定全局模型和训练过程的超参数,如全局通信次数、学习率等。然后,服务器将初始化的全局模型 M_{G_0} 和任务广播给终端节点。

2) 本地模型训练和更新:基于全局模型 M_G , t 表示当前迭代索引,每个终端节点分别使用其本地数据和设备更新本地模型参数 M_i ,在迭代数 t 下的参与者 i 的目标是找到最优参数 M_i 以实现最小化损失函数 $L(M_i)$,更新后的本地模型参数随后被发送到服务器。

3) 全局模型聚合和更新:服务器聚合来自参与者的局部模型,然后将更新后的全局模型参数 M_{i+1} 发送回终端节点。

4) 服务器希望最小化全局损失函数 $L(M_G)$,重复步骤 2) 和 3),直到全局损失函数收敛或达到预定训练精度。

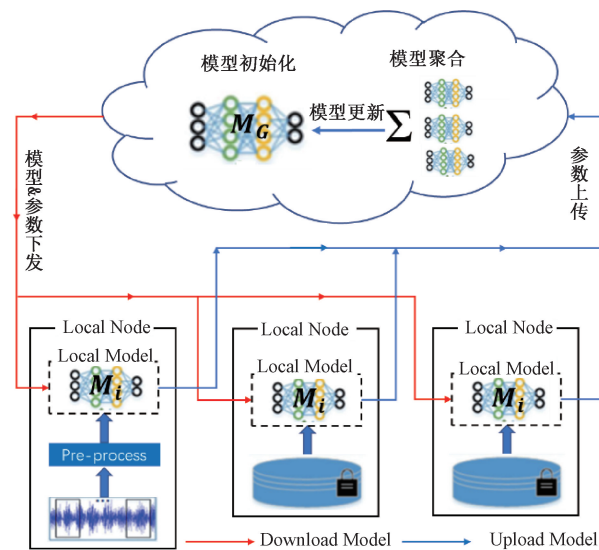


图 1 基于联邦学习的故障诊断基础框架

Fig. 1 Basic framework of fault diagnosis based on federated learning

在基于联邦学习的故障诊断流程中,局部模型训练和全局模型聚合对最终故障诊断任务起着决定性作用。对于局部模型,节点数据对模型性能影响巨大,若对节点数据的特征提取力度不足,则会降低终端节点模型训练质量;对于全局模型,多终端节点间的数据分布差异会导

致聚合后的全局模型分类精度较差。因此,本文以上述联邦学习故障诊断基础框架为基础,对终端节点进行特征提取增强并改善全局模型聚合算法,以综合化提升全局模型性能。

2 多小波系数增强动态聚合联邦深度网络诊断框架

2.1 终端节点多小波系数融合的特征增强

1) 基于小波包分解的数据特征表示

小波包分解^[15]是从小波分析延伸出来的一种对信号更加细致的分析与重构方法,它将频带进行多层次划分,对多分辨分析没有细分的高频部分进一步分解,并能够根据被分析信号的特征自适应地选择相应频带,使之与信号频谱相匹配,在时域和频域内都有表征信号局部信息的能力。

当原始输入振动信号为 $x(t)$ 时,小波包分解将信号分解为若干组小波系数,其中每组小波系数代表某一频段内的信息,图 2 为 3 层小波包分解示意图,其中右下标前一个数字为分解层数,右上标后一个数字为该层中频带编号。

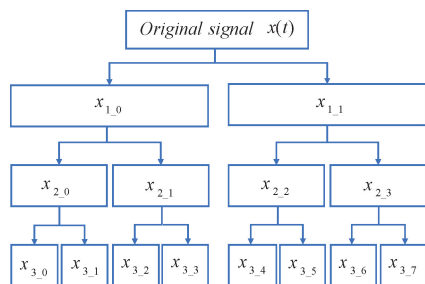


图 2 3 层小波包分解示意图

Fig. 2 Diagram for the 3-layer wavelet packet decomposition

小波包分解将原信号转换到 2^j (j 为分解层数) 个频带中,通过双尺度方程构造小波包,双尺度方程可以写作:

$$\begin{cases} \phi_{j,n}^{2i}(t) = \sum_k h_k \phi_{j-1,2n-k}^i(t) \\ \phi_{j,n}^{2i+1}(t) = \sum_k g_k \phi_{j-1,2n-k}^i(t) \end{cases} \quad (1)$$

式中: $\phi(t)$ 为母小波, j 为分解层数, n 为该层中节点编号, k 为位移因子。

第 j 层第 k 个频带的小波包系数可通过式(2)计算。可提取小波系数矩阵中的值重构信号。

$$\begin{cases} d_j^{2i}(n) = \sum_k h_k d_{j-1}^i(2n-k) \\ d_j^{2i+1}(n) = \sum_k g_k d_{j-1}^i(2n-k) \end{cases} \quad (2)$$

式中: $d(n)$ 为小波包系数, h_k 表示高通滤波器, g_k 表示低通滤波器。

基于小波包分解的故障诊断依赖于小波基函数以及对所得分解系数的选择。一方面,通常难以确定何种小波基函数与被分析信号最为匹配、以及哪些子带信号包含与机械健康状态(即正常和不同类型故障)最为相关的内在信息;另一方面,由于运行工况(如转速等)的变化往往会导致振动信号及其特征频带等的变化,一般来说,在不同的运行工况下诊断不同类型故障的最佳小波是不同的,同样,对于多种故障类型(例如,轴承内圈故障、外圈故障、滚动体故障等),某一种特定小波往往也不是最优选择,也即固定且单一的小波基函数、以及特定分解系数的选择方式在实际应用中存在局限。

2) 多小波系数融合的特征增强

多小波系数融合旨在提供对原信号更为全面的多样性描述,基于不同小波基函数的小波包分解所得系数给出了原信号不同角度的特征描述,从而提取适用于多工况下故障诊断的信号特征,进而提高对涉及多类型故障的诊断性能。

本文将不同小波基函数的小波包分解生成不同二维

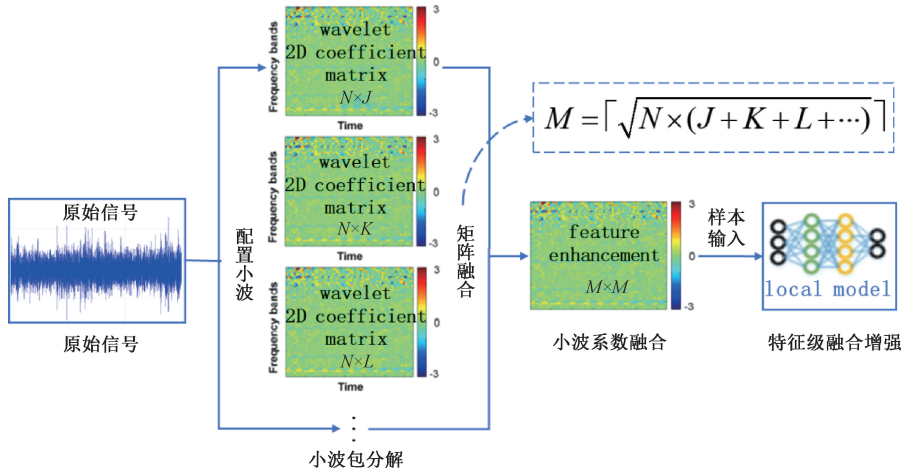


图3 多小波系数融合示意图

Fig. 3 Diagram of multiple wavelet coefficients fusion

2.2 聚合节点平衡模型贡献度的动态加权聚合

1) 非独立同分布对联邦聚合算法影响

在实际工业场景中,各节点拥有不同的工况导致局部数据与全局数据的分布不同,该特点在联邦学习场景下称为数据的非独立同分布(non-independent and identically distributed, Non-IID)问题^[16]。

联邦平均(federal averaging, FedAvg)是联邦学习的一种聚合算法^[17],终端节点在保护用户隐私的前提下进行联邦学习。FedAvg算法对收集到的本地模型进行加权平均,然后得到一个新的全局模型。聚合后的全局模

小波系数矩阵的能力融入到联邦学习模型训练中,即把多个不同小波基函数生成的二维系数矩阵进行组合,构造出一个新的多小波系数融合矩阵,并将其作为联邦学习终端节点模型的输入完成多小波包系数的特征级融合,从而提取更具判别性故障特征。

多小波系数融合示意如图3所示。首先配置一组小波基函数,分别利用不同的小波基函数对原始信号做 x 层小波包分解;然后将各小波包分解后得到的多个小波系数矩阵进行组合实现特征多样性增强;最后使用增强后的信号特征进行故障诊断模型训练。多小波系数融合矩阵的维度大小设置为:

$$M = \lceil \sqrt{N \times (J + K + L + \dots)} \rceil \quad (3)$$

式中: M 为融合后的多小波系数矩阵大小, $\lceil \cdot \rceil$ 表示向上取整, N 为单小波系数矩阵的长, J, K, L 分别表示不同的单小波系数矩阵的宽。

多小波系数融合通过组合同一信号基于不同角度的小波包分解特征信息,不同小波基从不同侧面给出数据的特征描述,提供了一种利用经典振动信号分析处理算法增强小样本深度学习的数据增强方法,降低因单终端节点小样本对局部模型训练的影响。

型参数 $M(w)$ 以广播的形式发送到每个参与训练的节点。假设有 P 个节点参与训练,那么在 t 轮的全局模型更新过程如式(4)所示:

$$\min_w \left\{ M(w) = \sum_{i=1}^P \frac{n_i}{N} M(w)_i \right\} \quad (4)$$

式中: P 为参与训练的节点数量, n_i 为第 i 个终端节点的训练样本数量, N 为所有参与训练的节点样本总数, $M(w)_i$ 为第 i 个节点在其本地数据集上训练的本地目标函数, n_i 和 N 构成了在聚合时第 i 个节点的权重。

常规基于均衡样本假设的 FedAvg 算法仅依赖于节

点间总数据量的差异、未考虑到各节点内样本特征空间不同带来的模型差异问题,导致其仅在数据分布相似的终端节点上具有较好的效果。

在非独立同分布数据场景下,联邦局部模型中部分工况缺失导致局部模型间存在性能差异,而为了识别更多工况必须联合所有节点数据信息,因此在联邦聚合时不仅要利用所有局部模型参数而且不同性能节点应给予不同聚合权重,从而提升全局模型分类性能。此外,精度常被用来衡量模型性能,在改善传统 FedAvg 算法中,精度被添加到模型聚合的权重中,但分类器的性能并没有明显提高^[18],在数量极度不平衡甚至缺失的情况下并不适用,因此联邦模型聚合算法需要考虑更多的因素。

2) 联邦动态加权的局部模型聚合

针对各终端节点数据非独立同分布问题,本文在 FedAvg 的基础上设计了一种动态加权聚合算法,相对于常规 FedAvg 算法,动态加权聚合算法充分考虑局部模型的召回率、精确度以及与全局模型的相似度等因素,如图 4 所示。

局部模型动态加权聚合算法核心在于局部模型贡献度分数的计算。贡献度由局部模型的 $F1$ 分数和局部模型与全局模型的欧氏距离 (euclidean distance, ED) 组成^[19]。 $F1$ 分数能够同时考虑模型的真阴率和真阳率,因此 $F1$ 分数可用来判断局部模型的性能优劣,其计算方式如式(5)所示。

$$F1_{ij} = \frac{2 \times Recall_{ij} \times Precision_{ij}}{Recall_{ij} + Precision_{ij}}$$

$$F1_i = \frac{\sum_{j=1}^k W_{ij} F1_{ij}}{\sum_{j=1}^k W_{ij}} \quad (5)$$

式中: k 是节点 i 的类别总数, $Recall_{ij}$ 表示召回率, $Precision_{ij}$ 表示精确度, W_{ij} 表示第 j 类节点 i 的权值,可以通过手动调整其值来优化节点内部各类别倾斜过程。 $F1_i$ 为第 i 个终端节点的 $F1$ 分数。

考虑通过限制局部模型的更新相对于全局模型的偏差量,以减少数据非独立同分布对全局模型带来的不利影响。本文计算局部模型与全局模型间的欧氏距离用来衡量模型间的相似度,计算公式如式(6)所示。

$$ED_i = \|M_G - M_i\|_2 \quad (6)$$

式中: M_G 表示全局模型参数, M_i 表示第 i 个节点的局部模型参数。

然后综合利用 $F1$ 分数与欧氏距离两个指标计算局部模型的贡献度分数,所设计的综合指标如式(7)所示。

$$W_i = \frac{2 \times F1_i \times ED_i}{F1_i + ED_i} \quad (7)$$

通过模型贡献度分数得到更新后的全局模型参数,如式(8)所示:

$$M_{t+1} = \frac{\sum_{i=1}^P n_i W_i M_i^{t+1}}{\sum_{i=1}^P n_i W_i} \quad (8)$$

最后全局模型可被用于测试数据的分析以及实际故障诊断,依据全局模型诊断精度等完成模型评价。

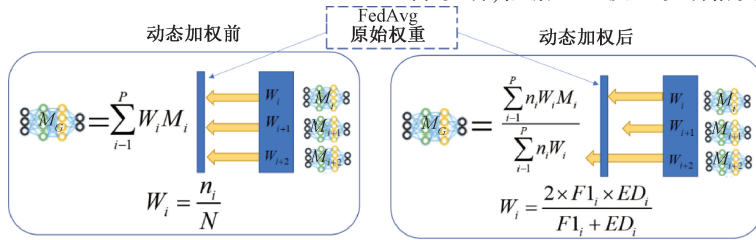


图 4 联邦动态加权聚合算法示意图

Fig. 4 Diagram for federated dynamic weighted aggregation algorithm

2.3 多小波系数增强动态聚合联邦深度网络

多小波系数增强动态聚合联邦深度网络整体框架如图 5 所示。该网络包括基于多小波系数融合的数据优化和基于动态加权聚合的联邦学习两个模块。在多小波系数融合模块中,可使用同类或多类小波基函数进行小波包分解;在联邦学习模块中,本文使用 CNN 模型作为联邦初始化模型。首先,通过不同小波基函数对原始信号做小波包分解;其次,利用多小波系数融合方法将得到的不同小波系数矩阵进行融合;之后,将多小波融合矩阵处理成 CNN 模型的输入进行模型训练;最后聚合节点利用

联邦动态加权聚合算法将局部模型的参数进行加权聚合得到全局模型进行故障诊断。

3 MWCE-FedDWA 的多工况故障诊断流程

本文所提小样本下多小波系数增强动态聚合联邦深度网络的多工况故障诊断流程如下:

1) 获取每个终端节点多工况的振动信号,选择不同小波基函数通过式(1)、(2)分别对其进行 x 层小波包分解,将不同小波基函数分解后的二维系数矩阵按照多小

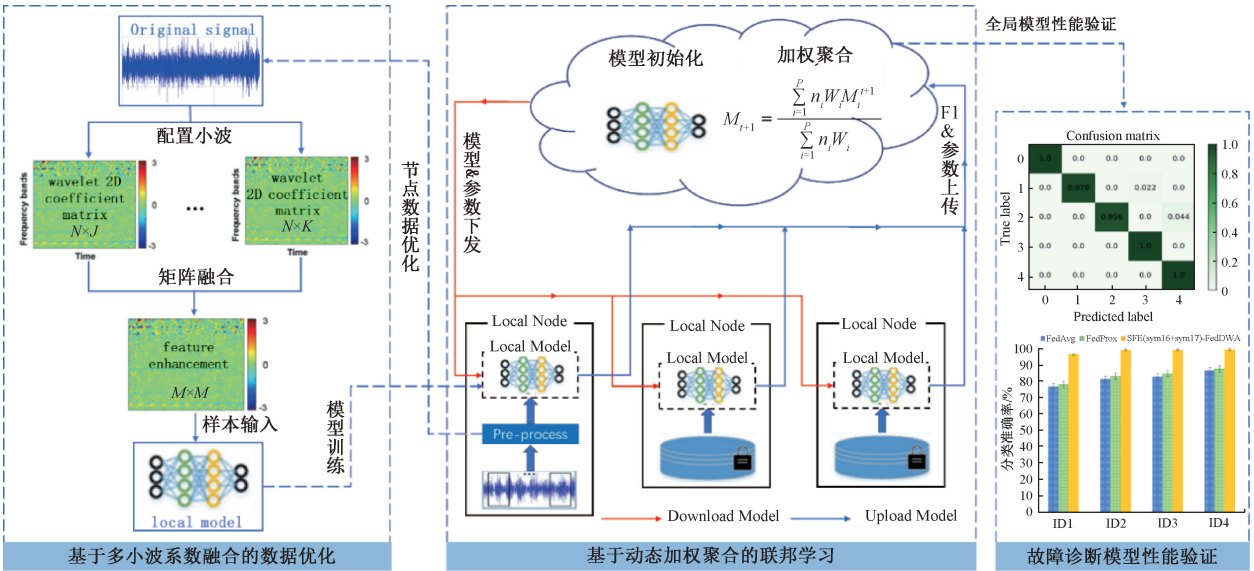


图5 多小波系数增强动态聚合联邦深度网络

Fig. 5 Multiple wavelet coefficients enhanced dynamic aggregation federal deep network

波系数融合的特征增强方法重构为 $M \times M \times 1$ 的样本,构建得到训练集 n_i 和测试集 t ,将训练集 n_i 输入终端节点模型中,将测试集 t 留在初始全局模型中用于最终分类测试;

2) 设置联邦学习框架下的初始化参数:batchsize,学习率,设置损失函数、终端节点模型迭代次数和全局通信次数等;

3) 将全局模型下发给终端节点得到局部模型,利用训练集 n_i 训练局部模型后并通过式(4)计算本地模型的F1分数,将局部模型参数与F1分数上传给受信任的第三方聚合节点;

4) 聚合节点利用上传的局部模型参数与拥有的全局模型参数通过式(5)计算欧氏距离 ED_i ,通过式(6)计算 W_i ,得到模型加权重;

5) 通过式(7)计算最终得到聚合后的模型参数进行全局模型更新,最后利用测试集 t 对全局模型进行模型评估;

6) 重复上述步骤3)~5),直到全局模型达到理想的分类精度,训练所得优化模型用于实际机械故障诊断任务。

4 实验验证

4.1 实验数据描述

为了验证所提出方法的有效性,本文利用动力传动系统诊断模拟器 DDS 进行故障模拟实验,分别采集了恒速多工况数据和变转速数据两种类型,图6展示了 DDS

主要结构。

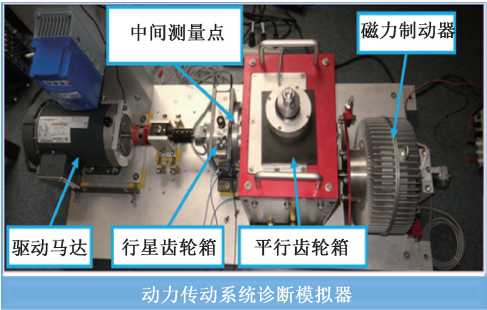


图6 实验测试台

Fig. 6 Experimental test rig

1) 恒速多工况数据描述

在 12.8 kHz 采样频率下,采集了 9 种不同工况(同一工况下转速和负载恒定)下的 5 种不同机械状态的振动信号,对于每个工况每种机械状态,原始振动信号被划分为 800 个样本,每个样本长度 1 024。具体实验样本组成以及工况具体信息如表 1、2 所示。注意,本文将工况编号处理是为了直观表示工况信息,但终端节点仅可获取样本类别标签,但无法获得其工况种类编号。

表1 实验样本组成

Table 1 Composition of experimental samples		
样本类型	类别标签	工况种类编号
正常	0	0,1,2,3,4,5,6,7,8
轴承复合故障	1	0,1,2,3,4,5,6,7,8
轴承滚动体故障	2	0,1,2,3,4,5,6,7,8
轴承内圈故障	3	0,1,2,3,4,5,6,7,8
轴承外圈故障	4	0,1,2,3,4,5,6,7,8

表 2 9 种工况具体信息

Table 2 Specific information of 9 working conditions

工况种类编号	电机转速/Hz	加载
0	20	0
1	20	4
2	20	8
3	30	0
4	30	4
5	30	8
6	40	0
7	40	4
8	40	8

本文模拟 4 个终端节点与 1 个聚合节点进行联邦学习故障诊断。为了模拟分布式小样本场景,对于每个工况每种机械状态,从 800 个样本中随机抽取共计 40 个样本,并平均分给各个终端节点(每个终端节点 10 个)作为训练样本。针对工况交叉数量(即 Non-IID 数据不平衡程度)对模型精度的影响,本文设置终端节点拥有不同工况训练样本的 4 个实验来验证,即每个终端节点含有全部 5 种样本类型但包含不同编号的工况种类信息,终端节点模型训练样本所包含的工况信息如表 3 中所示。

为了便于比较,4 个实验中的测试集样本数设置为一致,对于每个工况每种机械状态,从除训练样本外的剩余样本中随机抽取 20 个样本作为测试样本,即共计 900 个测试样本。

表 3 各节点中工况种类

Table 3 Types of working conditions for each node

实验 ID	工况种类数 K	节点 1	节点 2	节点 3	节点 4
1	3	0,1,2	2,3,4	4,5,6	6,7,8
2	4	0,1,2,3	2,3,4,5	4,5,6,7	0,1,7,8
3	5	0,1,2,3,4	2,3,4,5,6	0,5,6,7,8	0,1,2,7,8
4	6	0,1,2,3,4,5	1,2,3,4,5,6	0,4,5,6,7,8	0,1,2,6,7,8

2) 变转速数据描述

在 25.6 kHz 采样频率下,采集 5 种不同机械状态的变转速运行下振动信号,采样过程中电机先升速(20~40 Hz)再降速(由 40 Hz 降低),变速数据具体信息如表 4 所示。对于每种机械状态,原始振动信号被划分为 1 600 个样本,每个样本长度 1 024。

表 4 变速数据具体信息

Table 4 Variable speed data specific information

样本类型	类别标签	电机转速	加载
正常	0		
轴承复合故障	1	先升速(20~40 Hz)	
轴承滚动体故障	2	再降速(由	0
轴承内圈故障	3	40 Hz 降低)	
轴承外圈故障	4		

同样模拟 4 个终端节点与 1 个聚合节点进行联邦学习故障诊断。为了模拟分布式小样本场景,对于每种机械状态,从 1 600 个样本中随机抽取共计 40 个样本并平均分给各个终端节点(每个终端节点 10 个)作为训练样本,再将剩余的所有样本全部作为测试样本。每个终端节点含有全部 5 种样本类型,但各终端节点中样本转速信息并不一致。

4.2 实验参数设置

为了验证所提出方法的有效性,本节利用实验数据对比其他相关方法进行了实验验证。所有实验均在 Python 3.7.11 和框架 Pytorch-1.8.0+cu111 下使用 Intel(R) Core(TM) i5-10400F CPU、NVIDIA GeForce GTX 1080 和 12 GB RAM 进行。为了保证实验的公平性,所有方法在实验过程中都使用相同的模型网络架构和超参数。基本参数设置如表 5 所示。CNN 网络架构由 3 层卷积运算组成。每层卷积运算操作分别包括卷积层、最大池化层、批归一化层和激活层,激活层是使用线性修正函数,使用 Adam 优化器对网络进行更新,学习率为 0.01,训练批次大小为 32,本地模型训练 10 个 epoch,全局通信 50 次。为避免随机初始化对实验结果的波动,每组实验进行 5 次,取 5 次实验结果的最优值作为该组的最终实验结果。注意,为了实验简便,本文选择对原始信号进行 5 层小波包分解,同时实验中以两个小波基函数的融合为例开展终端节点特征增强,实际应用中可选择不同层小波包分解以及更多小波基函数作为特征增强。

表 5 MWCE-FedDWA 实验参数设置

Table 5 Parameters for the MWCE-FedDWA experiment

参数名称	参数值
卷积层数	3
池化层数	3
卷积核大小	3×1
池化尺寸	2
优化器	Adam
优化器学习率	0.01
批处理个数	32
本地迭代次数	10
全局迭代次数	50

4.3 恒速多工况数据结果分析

1) 对比实验与结果分析

为了证明所提算法的有效性,在 4 组不同工况交叉数量的小样本数据下,将使用 MWCE-FedDWA 算法所得结果与使用 Fedavg 算法和 Fedprox^[20] 算法所得结果进行对比分析,不同算法所得分类结果如图 7 所示。

如图 7 所示,MWCE-FedDWA 算法下训练得到的分类精度远远高于在 Fedavg 算法和 Fedprox 算法下的分类

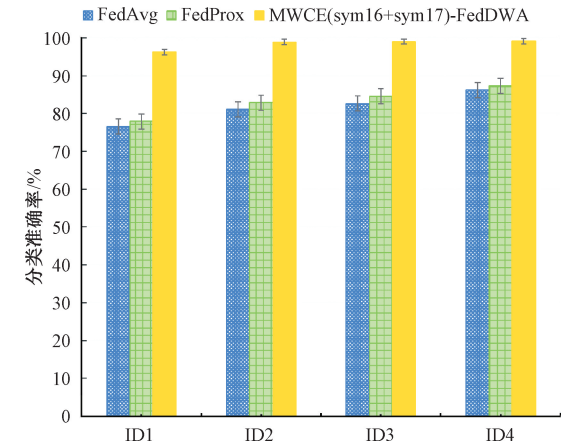


图7 不同算法分类结果对比

Fig. 7 Comparison of classification results for different methods

精度。如实验ID为1中,工况数据不平衡程度最强,模型在MWCE-FedDWA算法下的分类精度达到了96.2%,较对比算法提高了近20%,在实验ID为4时,所提算法的分类精度达99.1%,充分验证了本文所提方法适用于分布式小样本下的多工况故障诊断应用。

为了更进一步的观察本文所提算法在各个类别故障中的表现,绘制实验ID4对应的分类混淆矩阵,如图8所示,可以看出,标签0正常类别、以及标签3和标签4分别表示的内圈故障和外圈故障均可以被模型准确识别;标签1和2分别为复合故障和轴承滚动体故障,实际诊断任务中诊断难度相对较大,图8中所得准确率相较于其他几类略低,但仍取得了95%以上的准确率。

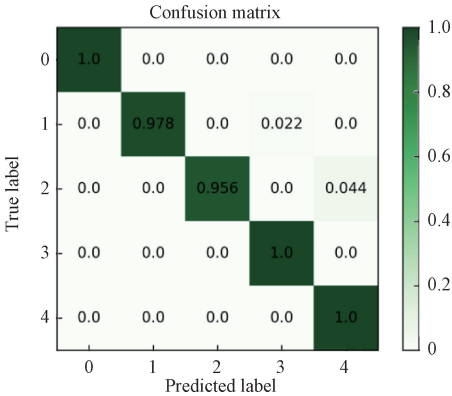


图8 ID4对应的分类混淆矩阵

Fig. 8 Classification confusion matrix corresponding to ID 4

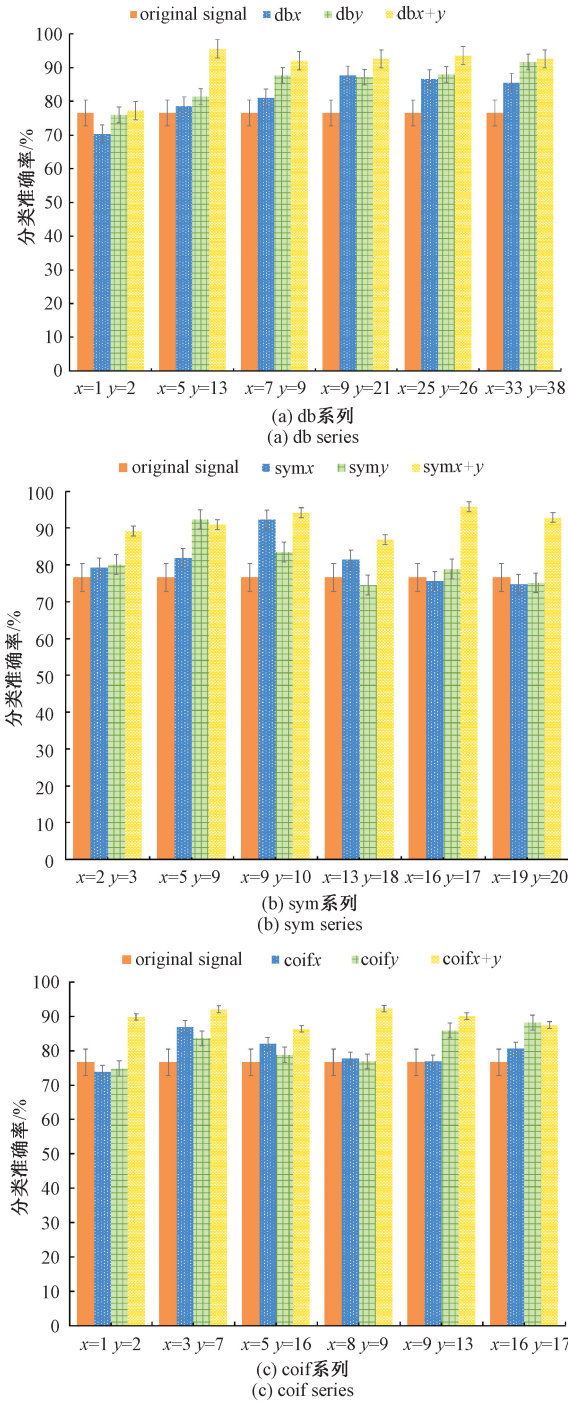
2) 恒速多工况消融实验以及泛化性验证

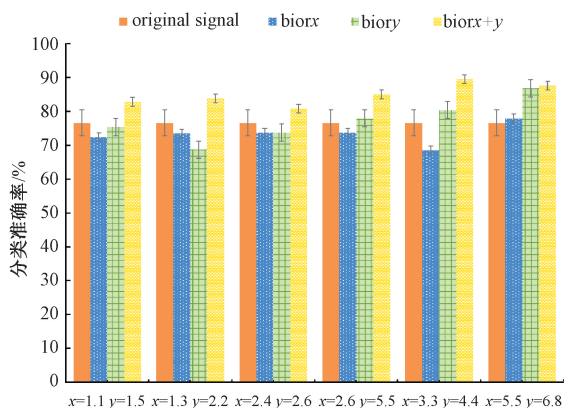
(1) 多小波系数融合的特征增强

有效性验证:对于多小波系数融合方法的有效性,在FedAvg算法的基础上进行验证。其中,局部模型中每种工况每类故障样本个数为10个,每个节点中包含全部5种类别但只包含3种工况,工况种类信息如表3实验

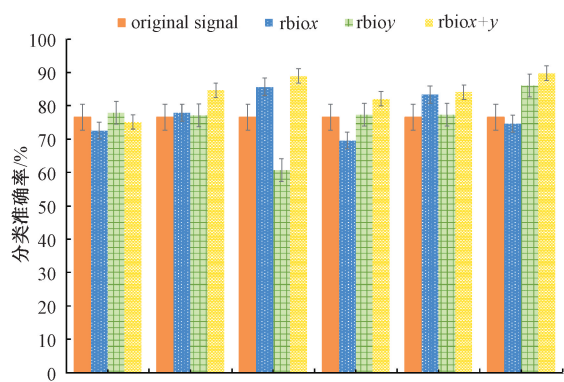
ID1。本次实验选取了5个系列的小波基函数(db系列、sym系列、coif系列、bior系列、rbio系列),共进行5组实验。实验结果如图9所示,第1列表示原始数据在FedAvg算法下的精度,第2、3列表示只使用单个小波包分解在FedAvg算法下的精度,第4列表示基于多小波系数融合方法的精度。

从图9实验结果可以得出,绝大多数情况下由多小波组成的数据来进行分类的精度高于原始数据和单个小波基函数处理的数据,小部分经过小波包分解后的数据分类精度低于原始数据是因为部分小波基函数中低





(d) bior 系列
(d) bior series



(e) rbio 系列
(e) rbio series

图 9 同系列混合小波基函数的实验结果

Fig. 9 Experimental results corresponding to mixed wavelet basis functions of the same series

阶滤波器的分解提取能力不强。而在小波包分解中,滤波器阶数相对越高,滤波器幅频特性的过滤带就越陡,越接近理想特征。图中,原始数据的分类精度为 76.6%,在单个小波包分解中当小波基函数为 sym9 时精度能达到 92.3%,而经过两个小波分解融合特征后即小波基函数为 sym16+sym17 时,分类精度可达到 95.8%,证明了所提出多小波系数融合的特征增强方法的有效性。

泛化性验证:为验证多小波系数融合方法的泛化性能,同样在 FedAvg 算法的基础上进行不同系列多小波实验验证。将上述 5 种系列小波混合,随机选取 6 组不同系列小波组合进行一次混合多小波实验。

由图 10 可以得到,从不同系列的小波基函数中进行小波组合仍能得到不错的效果,如 db9+sym9 组合后的分类精度可达到 94.3%,进一步验证了多小波系数融合特征增强的良好泛化能力。

(2) 联邦动态加权的局部模型聚合

为验证所提 FedDWA 算法在多工况 Non-IID 场景下的诊断以及泛化能力,将其实验结果与 FedAvg 和

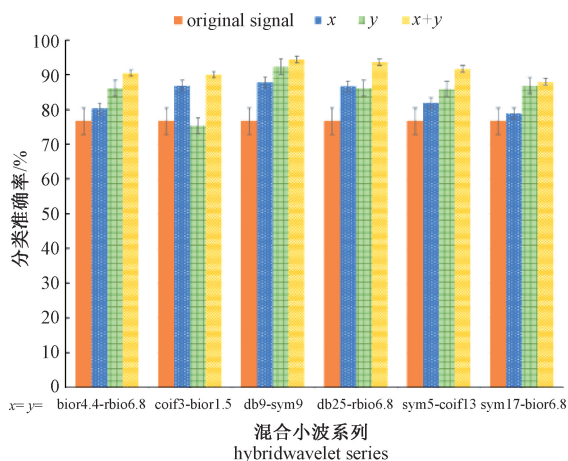


图 10 不同系列混合小波基函数实验结果

Fig. 10 Experimental results corresponding to different series of mixed wavelet basis functions

FedProx^[17] 两种典型算法所得结果进行比较,如图 11 所示。训练集不做数据优化处理,即使用原始数据对 4 种不同工况交叉的训练集进行训练,其中,局部模型中每种工况每类故障样本个数同样为 10,每个节点中工况种类信息如表 3 中 4 个实验 ID。

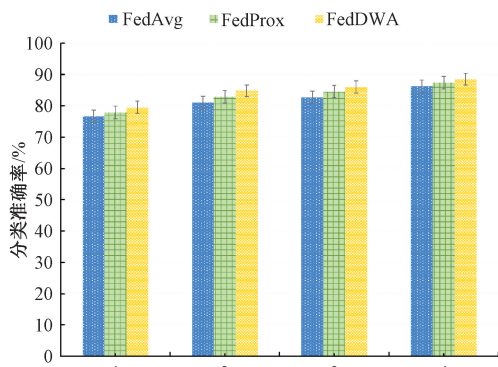


图 11 Non-IID 数据集上不同聚合算法结果对比

Fig. 11 Result comparison for different aggregation algorithms on Non-IID data sets

由图 11 知,联邦动态加权的局部模型聚合算法在 Non-IID 数据集下总体表现出了较大优势。考虑到节点含有的工况数量可一定程度反映 Non-IID 数据集的不平衡程度指标,图中实验 ID 为 1→4 时,节点含有的工况数量 K 分别为 3→6,显然,随着节点工况数量的增多,数据的不平衡程度减弱,模型分类精度逐渐增加,当 $K=6$ 时(实验 ID 为 4),FedDWA 算法精度达到 88.6%。总体而言,在 4 种不同程度的不平衡数据下,提出的 FedDWA 算法较对比算法均有明显优势,能够更好的支撑 Non-IID 数据集场景下故障诊断应用。

4.4 变转速数据结果分析

为进一步验证所提算法在变转速数据下的有效性,以随机选取的 3 组小波基函数为例开展变转速下故障诊断,实验结果如表 6 所示。

表 6 变转速下实验验证结果
Table 6 Experimental results of
variable speed were verified (%)

特征增强方法	模型聚合算法		
	FedAvg	FedProx	FedDWA
无特征增强(原始数据)	62.15	63.00	63.35
MWCE(db5+db13)	93.78	94.34	94.64
MWCE(db9+sym9)	89.89	90.59	91.10
MWCE(sym16+sym17)	91.99	92.51	93.31

从表 6 可以得到,经 MWCE 特征增强后再进行联邦故障诊断时,在 3 组不同小波基函数组合下,所提的 FedDWA 算法均得到优于对比聚合算法 FedAvg 和 FedProx 的诊断精度,验证了所提联邦动态加权聚合算法的有效性。同时,经过多小波系数增强后的联邦故障诊断精度显著优于无特征增强的原始数据下联邦故障诊断精度,验证了多小波系数融合特征增强方法在变转速下用于故障诊断的有效性。此外,值得注意的是,在 3 组不同小波基函数组合下,本文所提的多小波系数增强动态聚合联邦深度网络在变转速下均获得了 90% 以上的诊断精度。

5 结 论

本文提出了一种多小波系数增强动态聚合联邦深度网络用于分布式小样本下的多工况故障诊断,并进行试验验证,得到以下结论:

- 1) 将小波包分解与联邦深度网络相结合,提出多小波系数嵌入局部模型的特征级融合增强方法,节点端提取本地样本的多样化的多小波系数并组合嵌入局部深度模型完成特征级融合,以获取更具判别性故障特征,为小样本下特征提取提供新思路;
- 2) 提出一种联邦动态加权聚合算法用于全局联邦深度网络模型的构建,设计一种综合考虑局部模型 $F1$ 分数及其与全局模型偏离度的权重指标,以降低非独立同分布问题对聚合影响;
- 3) 所提方法中多小波系数融合特征增强和联邦动态加权聚合的联合运用可有效降低分布式节点工况分布不平衡对基于联邦深度网络的故障诊断的影响;与其他方法相比,在分布式场景下利用多个客户端的多种工况数据,训练得到能识别更多工况且诊断精度更高的模型;
- 4) 所提多小波系数增强动态聚合联邦深度网络在无

需数据共享前提下,利用分布式多节点端的小样本数据完成深度网络诊断模型构建和优化,适用于多节点协同、数据隐私保护需求下的机械健康监测与故障诊断应用。

参考文献

[1] BAROT A, KULKARNI P. Technological evolution in the fault diagnosis of rotating machinery: A review[J]. IOSR Journal of Mechanical and Civil Engineering, 2021(2): 18.

[2] 张焱,何姝钊,王平,等. 无转速计下变工况滚动轴承故障特征量化表征提取[J]. 仪器仪表学报, 2021, 41(8): 104-114.

ZHANG Y, HE SH B, WANG P, et al. Tacholeless quantitative characterization of rolling bearing fault feature under varying conditions[J]. Chinese Journal of Scientific Instrument, 2021, 41(8): 104-114.

[3] MA Y L, CHENG J SH, WANG P, et al. A novel Lanczos quaternion singular spectrum analysis method and its application to bevel gear fault diagnosis with multi-channel signals [J]. Mechanical Systems and Signal Processing, 2022, 168: 108679.

[4] MAURICIO A, GRYLLIAS K. Cyclostationary-based multiband envelope spectra extraction for bearing diagnostics: The combined improved envelope spectrum[J]. Mechanical Systems and Signal Processing, 2020, 149: 107150.

[5] HE F, YE Q. A bearing fault diagnosis method based on wavelet packet transform and convolutional neural network optimized by simulated annealing algorithm[J]. Sensors, 2022, 22(4): 1410.

[6] NI Q, JI J C, FENG K, et al. A fault information-guided variational mode decomposition (FIVMD) method for rolling element bearings diagnosis [J]. Mechanical Systems and Signal Processing, 2022, 164(5): 108216.

[7] MOSHREFZADEH A. Condition monitoring and intelligent diagnosis of rolling element bearings under constant/variable load and speed conditions [J]. Mechanical Systems and Signal Processing, 2021, 149: 107153.

[8] MORTIER R, HADDADI H, SERVIA S, et al. Distributed data analytics[J]. Privacy by Design for the Internet of Things: Building Accountability and Security, 2022, DOI: 10. 48550/arXiv. 2203. 14088.

[9] KONEN J, MCMAHAN H B, YU F X, et al. Federated learning: Strategies for improving communication efficiency[J]. NIPS Workshop on Private Multi-Party Machine Learning, 2016, DOI: 10. 48550/arXiv. 1610. 05492.

[10] ZHANG Z, XU X, GONG W, et al. Efficient federated

- convolutional neural network with information fusion for rolling bearing fault diagnosis[J]. *Control Engineering Practice*, 2021, DOI: 10. 1016/j. conengprac. 2021. 104913.
- [11] ZHOU F, ZHANG Z, LI S. Research on federated learning method for fault diagnosis in multiple working conditions[J]. *Complex Engineering Systems*, 2021, DOI:10. 20517/ces. 2021. 08.
- [12] ZHANG W, LI X, MA H, et al. Federated learning for machinery fault diagnosis with dynamic validation and self-supervision[J]. *Knowledge-Based Systems*, 2021, 213(1):106679.
- [13] 魏秀业,程海吉,贺妍,等. 基于特征融合与 ResNet 的行星齿轮箱故障诊断[J]. *电子测量与仪器学报*, 2022,36(5):213-222.
- WEI X Y, CHENG H J, HE Y, et al. Fault diagnosis of planetary gearboxes based on feature fusion and ResNet[J]. *Journal of Electronic Measurement and Instrumentation*, 2022,36(5):213-222.
- [14] 毛向向,王红军,韩凤霞,等. 基于深度卷积神经网络的机电系统故障分类识别方法[J]. *电子测量与仪器学报*,2021,35(2):87-93.
- MAO X X, WANG H J, HAN F X, et al. Fault classification and recognition of electromechanical system based on deep convolutional neural network[J]. *Journal of Electronic Measurement and Instrumentation*, 2021, 35(2):87-93.
- [15] KUMAR H S, SRINIVASA P P, VIJAY G S, et al. Wavelet transform for bearing condition monitoring and fault diagnosis: A review[J]. *International Journal of Comadem*, 2014, 17(1):9-23.
- [16] LI Q, DIAO Y, CHEN Q, et al. Federated learning on Non-IID data silos: An experimental study[C]. 2022 IEEE 38th International Conference on Data Engineering, Kuala Lumpur, Malaysia, 2022: 965-978.
- [17] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[C]. *International Conference on Artificial Intelligence and Statistics*, 2016.
- [18] FENG C, WANG Y, ZHAO Z, et al. Joint optimization of data sampling and user selection for federated learning in the mobile edge computing systems[C]. *International Conference on Communications Workshops*, 2020.
- [19] GOU L, ZHANG J, LI N, et al. Weighted assignment fusion algorithm of evidence conflict based on Euclidean distance and weighting strategy, and application in the wind turbine system [J]. *PloS One*, 2022, 17(1):e0262883.
- [20] LI T, SAHU A K, ZAHEER M, et al. Federated optimization in heterogeneous networks[J]. *Machine Learning and Systems*, 2018, DOI: 10.48550/arXiv.1812.06127.

作者简介



张焱(通信作者),2011年和2016年分别于重庆大学获得学士学位和博士学位,现为重庆邮电大学副教授,主要研究方向为机电装备智能诊断与安全服役、工业物联网。
E-mail: yanzhang@cqupt.edu.cn

Zhang Yan received his B. Sc. and Ph.D. degrees both from Chongqing University in 2011 and 2016, respectively. Now he is an associate professor at Chongqing University of Posts and Telecommunications. His main research interests include equipment intelligent diagnosis and safety management, and industrial internet of things.