

DOI: 10.19650/j.cnki.cjsi.J2209464

边缘计算生成式对抗网络差分进化滚动 轴承特征识别方法*

张效天, 王雪, 强振峰

(清华大学精密仪器系 北京 100084)

摘要:滚动轴承是旋转机械系统中保障安全运行重要组成部分之一。开展滚动轴承特征识别具有重要理论实际应用价值。通常采用的深度学习滚动轴承特征识别方法,需要有监督标记数据或无监督故障数据参与训练,标签和故障数据不易获取,无法满足滚动轴承特征识别需求。本文提出了一种边缘计算生成式对抗网络差分进化滚动轴承特征识别 EC-DE 法。该方法训练过程采用健康数据训练生成式对抗网络,通过学习健康数据分布规律进行滚动轴承健康特征识别。边缘端对比输入样本与生成式对抗网络生成样本差异性进行识别,根据输入样本健康置信度提前退出,提高系统实时性;云端采用差分进化算法搜索生成式对抗网络生成器输入潜空间,获得输入样本对应生成器输入潜变量,提高识别精度。本文方法在 CWRU 滚动轴承公共数据集上的识别正确率达 99.8% 且对超参数不敏感,推理阶段耗时降低,具有实际生产应用价值。

关键词:滚动轴承;边缘计算;生成式对抗网络;差分进化

中图分类号: TH133.3 TH165+.3 TH17 **文献标识码:** A **国家标准学科分类代码:** 510.40

An edge computing method for differential evolution of generative adversarial networks rolling bearing feature recognition

Zhang Xiaotian, Wang Xue, Qiang Zhenfeng

(Department of Precision Instrument, Tsinghua University, Beijing 100084, China)

Abstract: Rolling bearing is one of the most important components of rotating machinery systems to ensure safe operation. It is important to carry out studies on rolling bearing feature recognition for theoretical and practical application. The commonly used deep learning rolling bearing feature recognition methods require supervised labeled data or unsupervised fault data to participate in the training, and labels of data and fault data are not easily accessible to meet the rolling bearing feature recognition requirements. This article proposes an edge computing method for differential evolution of generative adversarial networks rolling bearing feature recognition, namely the EC-DE method. The training process uses only healthy data to train the generative adversarial networks and learn the distribution pattern of healthy data. The edge node compares the distribution difference between the input samples and the generative samples of generative adversarial networks for identification and exits early according to the health confidence level to improve the system's real-time performance. The cloud node uses a differential evolution algorithm to search the generator latent space of the generative adversarial networks to obtain the latent variables corresponding to the input samples, which improves the recognition accuracy. The proposed method achieves 99.8% accuracy on CWRU rolling bearing public data set and is insensitive to hyper-parameters, and the inference stage takes less time, which is valuable for a practical production application.

Keywords: rolling bearing; edge computing; generative adversarial networks; differential evolution

收稿日期: 2022-03-18 Received Date: 2022-03-18

* 基金项目: 国家重点研发计划支持项目(2018YFB2003200, 2018YFB2003500)、广东省重点领域研发计划项目(2019B010154001)、河北省创新能力提升计划项目(20540301D)资助

0 引言

滚动轴承是复杂旋转机械系统重要组成部分,轴承失效引起的机械设备系统问题占比超过50%^[1]。滚动轴承影响机械设备的运行安全性和可靠性,对其健康状态进行监测识别十分重要。

目前采用有监督深度学习方法,通过对标记训练数据进行学习和网络优化实现滚动轴承特征识别,其中卷积神经网络^[2](convolutional neural network, CNN)和长短期记忆神经网络^[3](long short-term memory, LSTM)是两种主流深度学习算法。Janssens等^[4]构建了一种卷积神经网络旋转机械特征学习系统,其性能明显优于采用人工设计特征和随机森林分类器的经典特征工程方法。Eren等^[5]提出一种用于轴承健康监测一维卷积神经网络自适应方法,故障识别正确率达到97%以上。宫文峰等^[6]改进CNN全连接层结构并与支持向量机结合,在滚动轴承故障识别实验中正确率高达99.86%。唐赛等^[7]实现了基于LSTM的端到端轴承故障识别方法,效果优于预先提取时频特征的LSTM方法。张昌凡等^[8]改进卷积双向门控循环神经网络实现良好滚动轴承故障识别性能。毕鹏远^[9]提出一种基于卷积运算和LSTM的滚动轴承诊断模型,相比于LSTM具有更高分类正确率。Chen等^[10]提出一种新型滚动轴承故障端到端智能诊断方法,采用多尺度CNN特征提取器和堆叠LSTM神经网络模型,较基于先验知识的传统智能诊断方法正确率更高。

有监督学习方法需要足够标记样本进行训练。实际应用场景中,标记样本数量较少,轴承故障数据难获取。因为轴承故障在初期难以识别,当故障发展到较为严重的阶段时振动数据已失真导致特征可用性降低,即使采集到故障数据,也存在样本数少,故障类别单一的局限,不能完整反映故障特征分布规律。如何在标记样本不足的情况下有效识别得到学者广泛研究。吕枫等^[11]采用少量标记样本训练深度关系网络,为未标记样本分配伪标签扩充标记样本,实现特征识别的少样本半监督方法。丁煦等^[12]提出一种基于半监督阶梯网络滚动轴承特征识别方法,采用生成式对抗网络拓展数据集,在少样本少标签条件下识别正确率达到95.5%。Xia等^[13]将模型分为两个阶段训练,先用未标记数据对去噪自动编码器进行无监督训练,再用少量标记数据进行有监督微调,充分利用了大量未标记数据,减小对数据标签依赖。虽然上述方法取得一定效果,但需要少量标记数据或未标记故障数据参与训练。仅健康数据参与训练的滚动轴承故障特征识别研究较少但已得到领域内学者的关注。Principi等^[14]基于深度自编码器较早使用健康数据实现

电机故障无监督识别并评估多种自编码器架构性能,实验结果表明基于多层感知机的自编码器识别精度优于基于CNN和LSTM自编码器。Brito等^[15]人工提取振动时频特征,采用最小协方差行列式^[16](minimum covariance determinant, MCD)、单类支持向量机^[17](one-class support vector machine, OCSVM)等多种传统机器学习算法实现旋转机械无监督故障识别,并对模型可解释性进行研究。

本文提出了一种健康数据参与训练的边缘计算(edge computing, EC)生成式对抗网络(generative adversarial networks, GAN)差分进化(differential evolution, DE)滚动轴承特征识别方法(EC-DE法),主要贡献是:1)采用差分进化计算和生成式对抗网络实现仅使用健康数据参与训练的滚动轴承故障特征识别,提高了识别精度和稳定性,精度对差分进化超参数不敏感,鲁棒性强;2)为上述滚动轴承特征识别方法设计了基于提前退出机制的边缘计算架构,提出EC-DE法,在保证识别精度前提下提高实时性;3)搭建边缘计算实验平台,在美国凯斯西储大学(CWRU)数据集上进行实验验证本文方法有效性。

1 特征识别边缘计算架构

边缘计算将部分计算任务部署在靠近轴承监测对象一侧^[18],根据设备计算资源和算法资源占用情况分配存储任务,EC-DE边缘计算方法改善特征识别推理阶段实时性。所述特征识别边缘计算系统采用提前退出机制,其硬件结构与数据流示意图如图1所示。包括监测对象、智能传感端、边缘端和云端4个部分,通过振动数据采集、预处理时频特征提取、边缘端无监督朴素法识别和云端无监督DE法识别实现边缘计算生成式对抗网络差分进化滚动轴承特征识别方法(EC-DE法)。

系统运行分为两个阶段,即训练准备阶段和推理识别阶段。训练准备阶段数据流如图1虚线所示,训练样本上传至云端训练模型,训练完成后云端将生成样本下放至边缘端。具体来说,推理识别前智能传感端需采集健康数据上传至云端训练,云端在健康训练样本上对抗式训练生成器与鉴别器。训练完成后云端生成器生成一定数量的生成样本下放至边缘端存储,即可开始正式推理识别。推理识别阶段数据流如图1实箭头所示。数据流起点是监测对象,即边缘计算系统进行实时监测的滚动轴承旋转机械对象。监测对象振动信号由智能传感端采集。智能传感端主要包括具有边缘计算能力的智能传感器和预处理单元,负责采集振动信号,进行预处理时频特征提取,并将特征样本上传至边缘端。

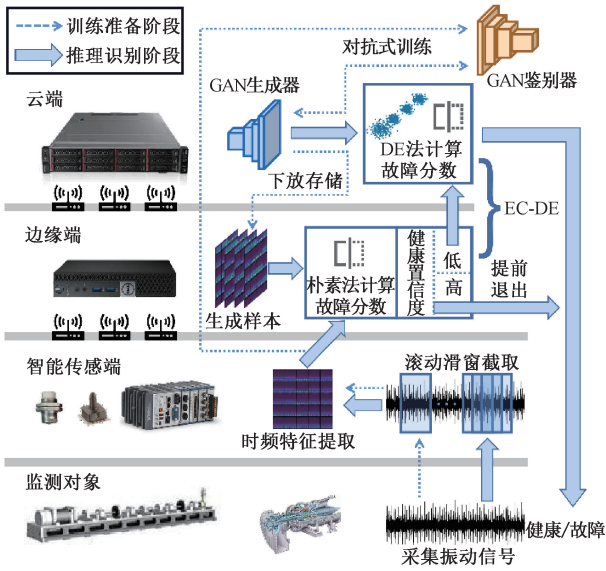


图1 特征识别边缘计算系统硬件结构与数据流图

Fig. 1 Feature recognition edge computing system hardware structure and data flow schematic

边缘端对输入样本进行初级朴素法识别。边缘端在计算资源和存储资源有限的情况下存储了经训练后的生成式对抗网络生成器生成的样本。生成样本近似服从真实健康样本分布。通过比较输入样本与存储的生成样本计算两者差异性可初步得到输入样本的健康置信度,根据置信度决定样本是否提前退出。

经边缘端初步朴素法识别计算输入样本健康置信度,置信度高的样本提前退出,置信度低的样本通过工业网络上传至云端进行深度DE法识别,得到精度更高的识别结果。在初级朴素法识别基础上,云端对输入样本采用差分进化算法在生成器输入空间中搜索输入潜变量 z ,最小化生成器输出 $G(z)$ 与输入样本的差异,提高特征识别精度。

2 生成式对抗网络特征识别方法

本节介绍滚动轴承振动数据采集后预处理与GAN网络训练,边缘端朴素法、云端DE法,以及基于提前退出机制的EC-DE法。

2.1 智能传感端预处理与云端生成式对抗网络训练

智能传感端采集轴承原始振动信号,采用连续小波变换提取时频特征。在轴承振动数据中依次截取4 096个数据点进行小波尺度为32的一维连续小波变换,再进行重采样尺度变换得到大小为 32×32 的特征图 \hat{x} ,最后对 \hat{x} 进行如式(1)所示的标准化得到样本 x ,使样本取值范围与生成式对抗网络生成器输出层取值范围对应,式中 $\max(\cdot)$ 和 $\min(\cdot)$ 分别表示求变量元素的最大

值和最小值。预处理环节在边缘计算系统中的智能传感端实现。

$$x = \frac{\hat{x} - \min(\hat{x})}{\max(\hat{x}) - \min(\hat{x})} - 0.5 \quad (1)$$

系统采用GAN模型进行无监督特征识别,采集健康数据样本作为训练集在云端进行训练。GAN作为无监督深度学习模型^[19],训练时无需获取数据标签,包括生成器和鉴别器两部分。本文仅使用轴承健康数据作为训练数据,同时训练生成器和鉴别器,采用Wasserstein GAN^[20]的优化目标如式(2)所示:

$$\min_{\theta_G} \max_{\theta_D} \mathbb{E}_{x \sim p_h} D(x) - \mathbb{E}_{z \sim p_z} D(G(z)) \quad (2)$$

式中 θ_G 和 θ_D 分别表示生成器和鉴别器的网络参数。式(2)代表了生成器和鉴别器两个相互对抗的目标函数。生成器目标是最大化鉴别器在生成样本上输出,生成尽量与训练健康样本相似的样本,输入潜变量 z ,将随机噪声分布 p_z 变换为接近健康训练数据服从的分布 p_h 。设 $p_z \sim N(0, I)$,即 p_z 服从高维高斯分布, I 为 $n_z \times n_z$ 单位矩阵,其中 n_z 为生成器潜变量 z 的维度。鉴别器目标是最大化鉴别器在训练样本和生成样本输出上的差异,区分训练样本和生成样本。训练完成后,生成器可生成与健康数据特征相似的样本。

2.2 边缘端朴素生成式对抗网络滚动轴承特征识别

边缘端朴素生成式对抗网络滚动轴承特征识别方法简称朴素法,配置于边缘端,其基本原理是经过生成式对抗网络生成器和鉴别器相互博弈的训练后,生成器能生成具有健康数据特征的样本,比较输入样本与生成样本的差异可初步反映输入样本的故障程度。当输入样本与生成样本差异较大,说明输入样本服从健康样本分布可能性小;反之当输入样本与生成样本差异较小,说明输入样本服从健康样本分布可能性大。为减小生成样本随机性对识别结果影响,计算故障分数时取多个生成样本进行平均处理,如式(3)所示。故障分数反映了输入样本故障程度,故障分数越大,故障程度越大。 N_{gen} 为初级朴素法识别中生成样本数。

$$A_{inf}(x_{test}) = \frac{1}{N_{gen}} \sum_{i=1}^{N_{gen}} \|x_{test} - G(z_i)\|_2 \quad (3)$$

边缘端计算资源有限,为简化计算流程,减小资源占用,边缘端进行初级朴素法识别时固定噪声 z_i ,边缘端只需预存储 N_{gen} 张二维时频特征图 $G(z_i)$ 并进行简单的矩阵运算可得到初步朴素法故障分数。预存储的特征图在模型训练完成后由云端下放到边缘端。

2.3 云端生成式对抗网络差分进化滚动轴承特征识别

云端生成式对抗网络差分进化滚动轴承特征识别方法简称DE法,配置于云端,其克服了边缘端初级朴素法简单地将输入样本与生成样本进行比较而忽略了生成样

本随机性影响特征识别结果的问题。为了充分挖掘生成器利用学习到的健康数据分布规律进行特征识别的潜力,减小生成样本随机性和输入样本中与健康状态无关特征形态对结果的影响,云端采用差分进化计算^[21]在生成器输入潜空间中搜索潜变量 \mathbf{z}_{test} ,使 $G(\mathbf{z}_{test})$ 与 \mathbf{x}_{test} 的差异尽量小^[22],得到深度 DE 法故障分数如式(4)所示。

$$A_{sup}(\mathbf{x}_{test}) = \min_{\mathbf{z}_{test}} \| G(\mathbf{z}_{test}) - \mathbf{x}_{test} \|_2 \quad (4)$$

差分进化算法是基于种群启发式搜索的全局优化算法,每次迭代计算具有内在并行性,可同时计算种群所有个体的目标函数值通过个体间合作与竞争协同搜索,参数自适应性好,适合在生成器潜空间进行变量搜索优化。差分进化算法包含初始化、变异、交叉和选择 4 个步骤。初始化:采用与 p_2 相同的高维高斯分布 $N(0, \mathbf{I})$ 生成特定种群数量的群体;变异:随机选择两个个体向量进行差分运算再乘以随机权重加在变异个体向量上;交叉:随机选择两个个体向量按照一定规则混合产生新个体;选择:计算当前种群的目标函数,筛选个体进入下一代,目标函数如式(5)所示。

$$f_{DE}(s) = \| G(s) - \mathbf{x}_{test} \|_2 \quad (5)$$

式中: s 表示种群个体向量, \mathbf{x}_{test} 表示输入测试样本。

2.4 推理识别阶段提前退出机制边缘计算架构

推理识别阶段,云端特征识别采用差分进化计算需反复进行生成器前向传播,运算量较大,在系统运行过程中产生一定延迟。为降低方法反应时间,将朴素法和 DE 法分别配置于边缘端和云端,基于提前退出机制设计了 EC-DE 边缘计算识别方法。

特征识别推理阶段分为 3 部分在三端进行,即智能传感端、边缘端和云端。

智能传感端对原始振动信号进行预处理,采用连续小波变换提取待查询振动波段的时频特征图。预处理与特征提取保证上传至边缘端数据大小格式统一,使样本与 GAN 网络生成器输出层和鉴别器输入层维度匹配。智能传感端位于数据源前端,进行数据特征提取可减小后端系统计算量,减小数据通讯传输量。

边缘端计算初级朴素法故障分数 $A_{inf}(\mathbf{x}_{test})$ 。设训练集上初级故障分数的中位数为 β 。若 $A_{inf}(\mathbf{x}_{test}) < \beta$,则故障识别结果为健康并提前退出推理识别。若 $A_{inf}(\mathbf{x}_{test}) \geq \beta$,将特征图上传至云端,进行进一步计算与识别。

云端对输入样本采用差分进化计算,得到最终深度 DE 法故障分数 $A_{sup}(\mathbf{x}_{test})$,若 $A_{sup}(\mathbf{x}_{test}) > \max(A_{sup}^{train})$,识别结果为故障。若 $A_{sup}(\mathbf{x}_{test}) \leq \max(A_{sup}^{train})$,识别结果为健康。 A_{sup}^{train} 为训练样本上的 A_{sup} 故障分数序列。

边缘计算架构根据不同计算单元计算能力分配不同计算任务,将计算向前端迁移,一方面缓解了云端计算压力,减小数据传输量,降低网络通信成本,另一方面改善

了系统实时性和精度之间的关系。提前退出机制的 EC-DE 边缘计算架构在样本健康置信度高的情况下可由边缘端直接得到识别结果,减少了云端传输与计算过程,在不影响特征识别精度情况下提高了系统反应实时性。云端只在边缘端判断当前样本健康置信度低情况下进行进一步计算。EC-DE 边缘计算架构改善了系统实时性与特征识别精度。

3 生成式对抗网络模型结构

生成式对抗网络分为生成器和鉴别器两个模型,网络结构如图 2 所示。训练阶段使用健康数据训练生成器和鉴别器,推理识别阶段采用生成器学习健康数据分布规律识别故障数据。

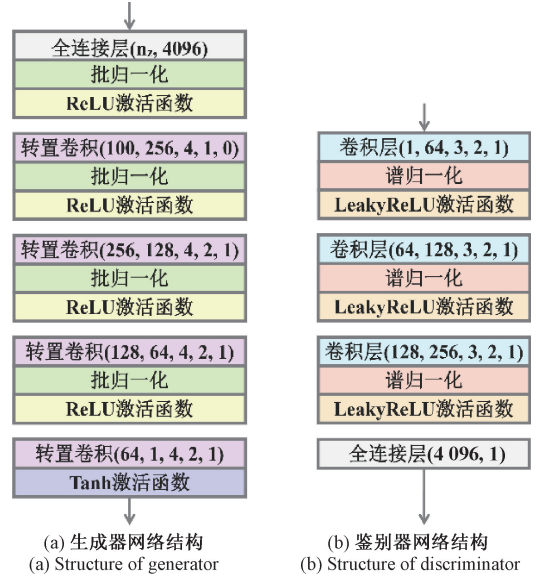


图 2 生成式对抗网络模型结构

Fig. 2 GAN model architecture

生成器采用深度卷积生成式对抗网络^[23]结构,采用多个转置卷积层扩大二维变量尺寸,实现上采样效果。图 2(a)展示生成器模型网络结构和参数,图中转置卷积层 5 个参数分别表示输入通道数、输出通道数、卷积核大小、卷积滑动步长和边界零填充宽度。为匹配特征识别任务,便于调整生成器输入潜空间维度,网络第一层设置为全连接层,目的是将生成器输入潜变量转换为 4 096 维向量,再变换为 256 个通道,尺寸为 4×4 的中间特征。生成器剩余部分为 3 层转置卷积层,依次将特征尺寸扩大为 8×8 、 16×16 和 32×32 ,同时将特征的通道数依次减小为 128、64 和 1。除了最后一层转置卷积层外,其余转置卷积层后面连接批归一化层和 ReLU 激活层,有助于改善模型训练收敛性,抑制模型坍塌等不稳定现象^[23]。最后一层转置卷积层接 Tanh 激活层,最终输出尺寸为

32×32的二维特征图,完成生成器从 n_z 维输入潜变量 z 到与样本尺寸相同二维特征图变换。

鉴别器采用三层卷积层与一层全连接层经典CNN结构。图2(b)展示了鉴别器模型网络结构和参数,卷积层后5个参数分别表示输入通道数、输出通道数、卷积核大小、卷积步长,即卷积时输入变量扩大倍数、输入变量零填充减小宽度。3个卷积层依次将输入特征尺寸由32×32压缩为16×16、8×8和4×4,通道数由1扩大为64、128和256。最后全连接层将卷积高层特征变换为单通道1维输出变量。为提高生成式对抗网络训练稳定性,在每个卷积层和全连接层后连接谱归一化层。谱归一化层计算权值矩阵谱范数对各层权值进行归一化以满足利普希茨系数约束条件,增强GAN训练稳定性^[24]。在卷积层对应的谱归一化层后,连接LeakyReLU激活层,这是GAN网络相关研究中常见设置^[23,25]。为与Wasserstein GAN目标函数匹配,最后全连接层不接激活层直接输出。

4 实验与讨论

本节介绍本文方法对CWRU滚动轴承公共数据集进行特征识别实验的具体设置和结果,验证本文方法有效性。

4.1 实验设置

CWRU滚动轴承数据集^[26]实验平台与主要部件如图3所示,电机驱动端轴承基座处放置振动传感器,用于采集轴承振动信号,采样频率为12 kHz。实验通过对电机施加不同负载设置了4种工况,利用电火花加工方式在滚动轴承3个位置引入3种不同深度的9种故障。总的来说,CWRU数据集包含4种工况下10类健康状态(1类健康数据和9类故障数据)的轴承振动数据。

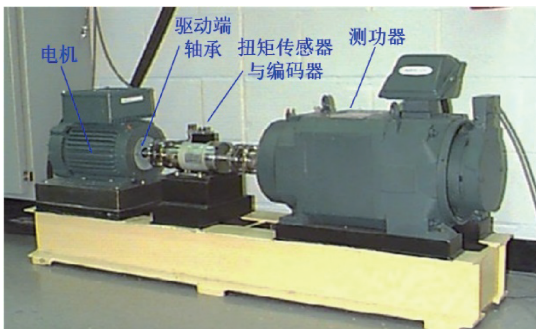


图3 CWRU滚动轴承数据集实验平台^[26]

Fig. 3 Experimental platform for CWRU rolling bearing dataset^[26]

采用滚动滑窗截取方式从每个工况健康数据中提取200个样本作为训练样本,900个样本作为测试样本,从

每个工况9类故障数据中分别提取100个样本作为精度测试集,分别提取10个样本作为耗时测试集。精度测试集的健康样本数和故障样本数相等,能更好地反映方法故障识别性能;耗时测试集的故障样本数为健康样本数的1/10,更符合实际应用场景中故障样本出现概率低的情况。故实验使用精度测试集得到方法精度指标结果,使用耗时测试集得到方法耗时指标结果。训练集和测试集信息如表1所示。

表1 训练集和测试集样本数量分布

Table 1 Number of samples in the training and test sets

数据集	健康样本数	故障样本数
训练集	200×4	0
精度测试集	900×4	100×9×4
耗时测试集	900×4	10×9×4

为构造边缘计算架构,在硬件实验平台方面本文分别使用MINI主机和高性能服务器作为边缘端和云端,硬件配置如表2所示。边缘端运行朴素生成式对抗网络滚动轴承特征识别方法。云端运行所提出的生成式对抗网络差分进化滚动轴承特征识别方法。

表2 边缘计算实验平台硬件配置

Table 2 Hardware configuration for edge computing experimental platform

平台	设备类型	CPU	内存/GB	独立显卡
边缘端	迷你主机	Core i7-9700T	16	无
云端	机架式服务器	Xeon 8280	256	Tesla A100

实验结果采用正确率、精确率和F1分数等文献中常用的指标对方法识别性能进行评价。为减小实验随机性对结果的影响,将每次实验重复10次进行平均得到最终实验结果。

4.2 本文方法性能结果

为获得实验结果,在健康数据上训练GAN模型。训练过程设置训练集batch size为16,采用RMSprop优化算法更新模型参数,初始学习率为0.0001,训练轮次为2000。训练完成后搭建边缘计算测试平台。实验分为3部分,分别测试朴素法、DE法和EC-DE边缘计算方法识别性能和推理识别阶段耗时情况。测试2.2节所述朴素法流程:在边缘端存储 N_{gen} 个生成样本采用朴素法对测试数据在本地识别。测试2.3节所述DE法流程:云端采用DE法对测试数据在本地识别。测试2.4节所述EC-DE边缘计算方法流程:将测试数据保存至边缘端,边缘端采用朴素法得到初步识别结果,决定测试样本是否提前退出,将健康置信度低的样本上传至云端采用DE

法识别,修正初步识别结果得到 EC-DE 边缘计算方法识别结果。设置朴素法中生成样本数 N_{gen} 为 10,进化计算的种群数量为 20,迭代次数为 10。

为与参考文献方法对比,选取文献[14]中精度最高的多层感知机自编码器方法、文献[15]中综合识别性能最好的 MCD 方法和常用的 OCSVM 方法与本文方法进行比较。为确保可比性,对比方法使用的训练集和测试集均与本文方法保持一致。文献[14]多层感知机自编码器方法只输出样本故障分数而不进行健康与故障识别,本研究中选取训练集样本中最高的故障分数作为阈值进行健康与故障识别得到相应精度指标。OCSVM 方法采用高斯核函数,超参数 γ 分别取 $2^{-1}, 2^{-2}, \dots, 2^{-6}$, 根据正确率选取最优 γ 。其他方法超参数设置与参考文献保持一致。

上述方法 10 次实验精度和耗时指标的均值和标准差如表 3 所示,其中耗时为推理识别阶段得到全部测试样本最终识别结果所用耗时,不包括模型训练耗时。表 3 中每项指标最优的两种方法对应指标加粗表示。本文边缘端朴素法、云端 DE 法,以及基于提前退出机制的 EC-DE 法 10 次实验故障识别正确率如图 4 所示。

表 3 特征识别方法性能结果

Table 3 Feature recognition method performance results

方法	正确率/%	精确率/%	F1 分数/%	耗时/s
MCD	95.60(0)	99.70(0)	95.80(0)	4.3(0)
OCSVM	96.40(0)	100.00(0)	96.50(0)	1.0(0)
文献[14]	63.40(0.02)	100.00(0)	73.23(0.01)	3.2(0.1)
朴素法	96.68(3.93)	93.38(7.86)	96.40(4.42)	0.1(0)
DE 法	99.87(0.16)	100.00(0)	99.87(0.16)	114.0(1.6)
EC-DE 法	99.80(0.24)	100.00(0)	99.80(0.24)	60.2(4.9)

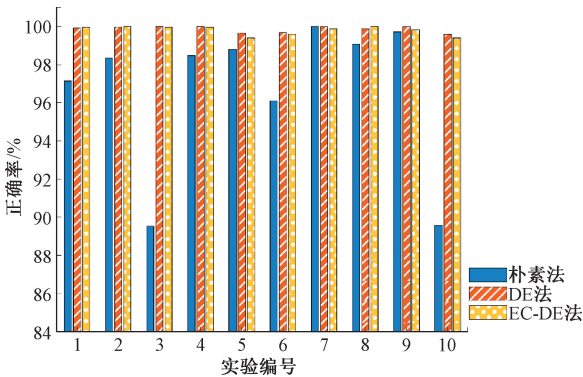


图 4 特征识别方法正确率实验结果

Fig. 4 Experimental results of the correctness of feature recognition methods

实验结果验证了所提出 DE 法与 EC-DE 法精度指标优于朴素法与其他对比方法,朴素法实时性优于其他方法。EC-DE 法的边缘计算架构在不影响精度性能情况下充分利用朴素法实时性优势,推理阶段耗时相比于 DE 法降低了 48.18%。

实验结果表明:EC-DE 法与 DE 法正确率接近,分别为 99.80% 和 99.87%,10 次实验结果标准差小,算法稳定性好,说明边缘计算架构不会影响 DE 法性能。理论上,只要边缘端在提前退出机制阶段将全部故障数据上传至云端分析,提前退出推理样本为健康样本,DE 法和边云协同法精度一致。实验过程,提前退出故障样本数为 0,说明提前退出样本全部为健康样本,EC-DE 识别性能不受影响,实验结果与理论分析一致。提前退出机制使 EC-DE 法耗时减小,在不损失精度的情况下提高了推理识别阶段实时性。朴素法识别性能波动大,稳定性差,因为生成器生成样本具有随机性,计算故障分数时没有考虑输入样本特征。其计算过程简单,资源占用小,实时性好,适合部署在边缘前端进行初步识别。

4.3 测试集故障样本比例与实时性分析

对测试集中故障样本比例与计算耗时关系开展研究,验证所提方法实时性,保持上述参数和测试集中的健康样本数不变,改变测试集故障样本数,得到 DE 法和 EC-DE 法推理识别阶段耗时与测试集故障样本数占比关系曲线如图 5 所示。图 5 中曲线数值为 10 次实验平均值,浅色阴影区域为对应方法 10 次实验最小值和最大值所包围区域。测试集健康样本数为 3 600,将测试集故障样本设置为健康样本数的 10%~100%,即故障样本数为 360~3 600。

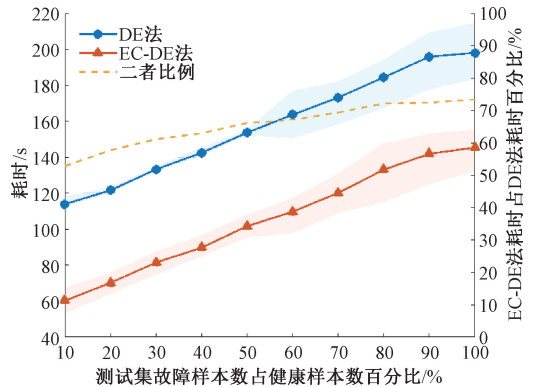


图 5 本文方法耗时情况与测试集故障样本数占比关系曲线

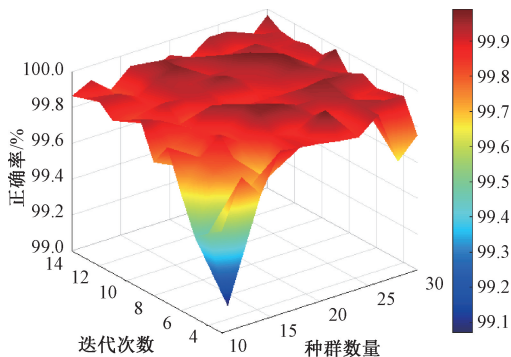
Fig. 5 The relationship between the time consumption of the methods and the number of fault samples in the test set

实验结果表明:故障样本数占数据集比例越小,EC-DE 法相比于 DE 法实时性优势越大。当故障样本数为健康样本数 10%,EC-DE 法相比于 DE 法耗时降低了

48.18%。因提前退出机制将健康置信度高样本提前退出推理阶段减小耗时,故障样本被上传至云端进一步计算。实际应用场景中,故障样本只占极少数,EC-DE法耗时显著降低,实际应用价值明显。

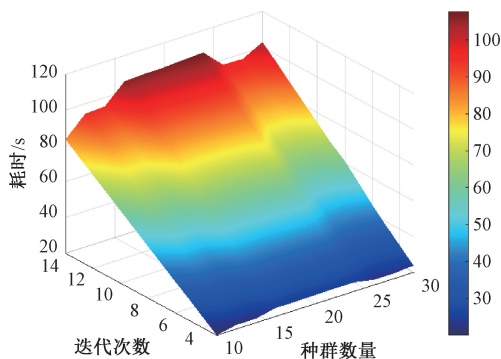
4.4 差分进化计算超参数与性能分析

对差分进化计算超参数与边云协同法正确率和实时性关系开展研究,设置差分进化迭代次数为3~15,种群数量为10~30,改变两个超参数得到对应正确率和耗时结果,绘制三维关系图如图6所示。



(a) 改变差分进化计算超参数EC-DE法正确率实验结果

(a) Experimental results of the correctness of EC-DE method with varying hyper-parameters of differential evolution



(b) 改变差分进化计算超参数EC-DE法耗时实验结果

(b) Experimental results of the time consumption of EC-DE method with varying hyper-parameters of differential evolution

图6 EC-DE法精度实时性指标同差分进化超参数关系图

Fig. 6 Correctness and time consumption of the EC-DE method in relation to hyper-parameters of differential evolution

滚动轴承特征识别正确率方面,通过改变进化计算迭代次数和种群数量,正确率指标稳定,均在99%以上。在迭代次数大于10,种群数量大于20情况下正确率指标稳定在99.7%以上,EC-DE法的特征识别能力对差分进化超参数不敏感,鲁棒性强。实时性方面,方法耗时随迭代次数成比例增加,随种群数量缓慢增加。因增大迭代次数会增加串行计算任务,导致耗时成比例增加。增大种群数量在差分进化计算中可转换为并行计算任务并利用云端服务器并行计算能力进行处理,从而耗时对种群数量超参数不敏感。

5 结 论

本文提出一种生成式对抗网络差分进化滚动轴承特征识别方法,训练过程只需健康数据,推理识别过程采用差分进化算法搜索输入样本对应生成器输入潜变量,提高特征识别精度和稳定性,识别结果对超参数不敏感,鲁棒性强。基于提前退出机制边缘计算架构提出EC-DC法,该方法匹配了不同计算资源与算法,分别将预处理特征提取、朴素生成式对抗网络滚动轴承特征识别方法和生成式对抗网络差分进化滚动轴承特征识别方法配置于智能传感端、边缘端和云端,在保证识别精度情况下提高了方法实时性。在CWRU滚动轴承数据集进行实验,结果表明EC-DE法与DE法精度相同,精度和稳定性得到提升,正确率高达99.8%,比朴素法高3.2%。同时EC-DE法耗时相比于DE法减小,在故障样本数为健康样本数10%情况下,耗时减小了48.2%,适合实际应用场景中故障样本少的情况。本文提出的边缘计算生成式对抗网络差分进化滚动轴承特征识别方法无需故障样本参与训练,解决了标签和故障数据获取问题,滚动轴承特征识别精度高,实时性好,具有重要理论和实际应用价值。

参考文献

- [1] NANDI S, TOLIYAT H A, LI X. Condition monitoring and fault diagnosis of electrical motors—A review [J]. IEEE Transactions on Energy Conversion, 2005, 20(4): 719-729.
- [2] LECUN Y, BOSER B, DENKER J, et al. Handwritten digit recognition with a back-propagation network [J]. Advances in Neural Information Processing Systems, 1990, 2: 396-404.
- [3] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
- [4] JANSSENS O, SLAVKOVIKJ V, VERVISCH B, et al. Convolutional neural network based fault detection for rotating machinery [J]. Journal of Sound and Vibration, 2016, 377: 331-345.
- [5] EREN L. Bearing fault detection by one-dimensional convolutional neural networks [J]. Mathematical Problems in Engineering, 2017: 1-9.
- [6] 宫文峰,陈辉,张美玲,等. 基于深度学习的电机轴承

- 微小故障智能诊断方法[J]. 仪器仪表学报, 2020, 41(1):195-205.
- GONG W F, CHEN H, ZHANG M L, et al. Intelligent diagnosis method for incipient fault of motor bearing based on deep learning [J]. Chinese Journal of Scientific Instrument, 2020,41(1):195-205.
- [7] 唐赛,何苻兮,张家悦,等. 基于长短期记忆网络的轴承故障识别[J]. 汽车工程学报,2018,8(4):297-303.
- TANG S, HE X X, ZHANG J Y, et al. Bearing fault identification based on long short-term memory networks[J]. Chinese Journal of Automotive Engineering, 2018,8(4):297-303.
- [8] 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]. Chinese Journal of Scientific Instrumentation, 2021, 35(11): 61-67.
- [9] 毕鹏远. 一种基于 Conv-LSTM 的滚动轴承故障诊断方法[J]. 机电工程技术,2021,50(11):113-115.
- BI P Y. A rolling bearing fault diagnosis method based on Conv-LSTM [J]. Mechanical & Electrical Engineering Technology,2021,50(11):113-115.
- [10] CHEN X, ZHANG B, GAO D. Bearing fault diagnosis base on multi-scale CNN and LSTM model [J]. Journal of Intelligent Manufacturing, 2021, 32(4): 971-987.
- [11] 吕枫,王义,阮胡林,等. 深度嵌入关系空间下齿轮箱标记样本扩充及其半监督故障诊断方法[J]. 仪器仪表学报,2021,42(2):55-65.
- LYU F, WANG Y, RUAN H L, et al. Labeled sample augmentation based on deep embedding relation space for semi-supervised fault diagnosis of gearbox [J]. Chinese Journal of Scientific Instrument,2021, 42(2):55-65.
- [12] 丁煦,王栋,翟华. 半监督阶梯网络和 GAN 在滚动轴承故障诊断的应用 [J]. 机械设计与制造,2022(5): 152-156.
- DING X, WANG D, ZHAI H. Application of semi-supervised ladder network and GAN in fault diagnosis of rolling bearing [J]. Machinery Design & Manufacture, 2022(5):152-156.
- [13] XIA M, LI T, LIU L, et al. Intelligent fault diagnosis approach with unsupervised feature learning by stacked denoising autoencoder[J]. IET Science, Measurement & Technology, 2017, 11(6): 687-695.
- [14] PRINCIPI E, ROSSETTI D, SQUARTINI S, et al. Unsupervised electric motor fault detection by using deep autoencoders[J]. IEEE/CAA Journal of Automatica Sinica, 2019, 6(2): 441-451.
- [15] BRITO L C, SUSTO G A, BRITO J N, et al. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery[J]. Mechanical Systems and Signal Processing, 2022, 163: 108105.
- [16] ROUSSEEUW P J, DRIESSEN K V. A fast algorithm for the minimum covariance determinant estimator [J]. Technometrics, 1999, 41(3): 212-223.
- [17] SCHÖLKOPF B, WILLIAMSON R C, SMOLA A, et al. Support vector method for novelty detection [J]. Advances in Neural Information Processing Systems, 2000, 12: 582-588.
- [18] SATYANARAYANAN M. The emergence of edge computing[J]. Computer, 2017, 50(1): 30-39.
- [19] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [J]. Advances in Neural Information Processing Systems, 2014: 2672-2680.
- [20] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks [C]. International Conference on Machine Learning, PMLR, 2017: 214-223.
- [21] STORN R, PRICE K. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces [J]. Journal of Global Optimization, 1997, 11(4): 341-359.
- [22] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery [C]. International Conference on Information Processing in Medical Imaging, Springer, Cham, 2017: 146-157.
- [23] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative

adversarial networks[J]. ArXiv Preprint, 2015, ArXiv: 1511.06434.

- [24] MIYATO T, KATAOKA T, KOYAMA M, et al. Spectral normalization for generative adversarial networks [J]. ArXiv Preprint, 2018, ArXiv: 1802.05957.
- [25] SHAO S, WANG P, YAN R. Generative adversarial networks for data augmentation in machine fault diagnosis[J]. Computers in Industry, 2019, 106: 85-93.
- [26] SMITH W A, RANDALL R B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study[J]. Mechanical Systems and Signal Processing, 2015, 64: 100-131.

作者简介



张效天, 2019 年于北京航空航天大学获得学士学位, 现为清华大学精密仪器系博士研究生, 主要研究方向为故障诊断、生成式对抗网络和边缘计算。

E-mail: zxt19@mails.tsinghua.edu.cn

Zhang Xiaotian received his B.Sc. degree from Beihang University in 2019. He is currently a Ph.D. candidate in the Department of Precision Instrument at Tsinghua University. His main research interests include fault diagnosis, generative adversarial networks and edge computing.



王雪(通信作者), 1994 年于华中科技大学获博士学位, 现为清华大学精密仪器系长聘教授, 博士生导师。主要研究方向为智能感知与精密测量。

E-mail: wangxue@mail.tsinghua.edu.com

Wang Xue (Corresponding author) received his Ph.D. degree from Huazhong University of Science and Technology in 1994. He is currently a tenured professor and a Ph.D. advisor in the Department of Precision Instrument at Tsinghua University. His main research interests include smart sensing and precision measurement.