

基于多分支 CNN 与改进级联森林的故障诊断

丁伟哲, 冉瑞生*, 胡子成

(重庆师范大学计算机与信息科学学院, 重庆 401331)

摘要:在故障诊断领域,卷积神经网络(CNN)、深度森林等深度学习模型因表现优异而备受关注。然而,单分支卷积神经网络提取的故障特征有限;深度森林的多粒度扫描针对不同的数据集需要重新设计和调整参数。本文提出了一种多分支 CNN 与改进的级联森林相结合的混合深度学习模型,首先设计了具有不同卷积核大小的多分支 CNN,可以并行提取多样特征,适应复杂任务需求。其次,由于极端梯度提升比随机森林更能处理非线性数据,将级联森林中的一个随机森林替换为极端梯度提升。同时,这种部分替换结合了不同算法的优势,优化了模型的整体性能。最后,将多分支 CNN 与改进的级联森林相结合,构建了一个混合深度学习模型。在 3 个轴承数据集和 1 个转子数据集上进行了实验,结果表明,所提出的模型在故障诊断中表现出很强的有效性。

关键词:故障诊断;卷积神经网络;级联森林;极端梯度提升

中图分类号: TP391

文献标志码: A

Fault diagnosis based on multi-branch CNN with improved cascade forest

DING Weizhe, RAN Ruisheng*, HU Zicheng

(College of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: In fault diagnosis, deep learning models such as Convolutional Neural Networks (CNN) and Deep Forest have demonstrated outstanding performance, attracting significant attention. However, single-branch CNNs extract limited fault features, and the multi-grained scanning in Deep Forest requires redesigning and adjusting parameters for different datasets. This paper proposes a hybrid deep learning model that combines a multi-branch CNN with an improved cascade forest. Firstly, a multi-branch CNN with different convolutional kernel sizes is designed to extract diverse features in parallel, meeting the needs of complex tasks. Secondly, since Extreme Gradient Boosting (XGBoost) handles nonlinear data better than random forest, one random forest in the cascade forest is replaced with XGBoost. This partial replacement leverages the advantages of different algorithms, optimizing the model's overall performance. Finally, a hybrid deep learning model combines the multi-branch CNN and the improved Cascade Forest. Experiments conducted on three bearing datasets and one rotor dataset demonstrate the proposed model's strong effectiveness in fault diagnosis.

Key words: fault diagnosis; CNN; cascade forest; XGBoost

对轴承和齿轮准确的故障诊断是保证设备安全运行的基础,对工业领域的持续发展至关重要^[1-3]。

随着技术的发展,利用计算机和人工智能诊断机械故障成为一个热门话题^[4]。故障诊断的传统方法是通过传感器采集旋转机械的一维振动信号,然后使用傅里叶变换(FFT)^[5]、奇异谱分析(SSA)^[6]、小波变换(WT)^[7]等信号处理方法提取数据特征,再将提取出的特征输入到支持向量机(SVM)^[8]、K近邻(KNN)^[9]等分类器中进行故障诊断。

2010年以来,不依赖专家经验的卷积神经网络(CNN)广泛应用于智能故障诊断并得到了高度的认可。例如,Wen等^[10]通过归一化振动数据转化为灰度图像,输入到改进的LeNet-5网络进行故障诊断;Islam等^[11]利用小波包变换(WPT)得到的缺陷深度比(DDR)值作为ADCNN的二维可视化输入数据;Wang等^[12]提出将3个注意力机制模块与1DCNN相结合,并应用于轴承故障诊断;Zhang等^[13]提出一种将原始信号转化为灰度图像的方法,将其作为输入数

收稿日期:2024-10-30

基金项目:重庆市教育委员会科学技术研究项目(KJZD-K202100505);重庆市技术创新与应用发展项目(cstc2020jcsx-msxmX0190,cstc2019jcsx-mbdxX0061)

作者简介:丁伟哲(1998—),男,硕士研究生,专业方向为智能故障诊断。

*通信作者:冉瑞生(1976—),男,教授,从事机器学习方向的研究,e-mail: rshran@cqnu.edu.cn。

据,使用具有 2 个 Dropout 层和 2 个全连接层的卷积神经网络进行故障诊断;Zhao 等^[14]通过使用软阈值作为非线性变换层,并自适应地确定子网络内的阈值,设计了一种适应故障噪声数据的残差网络。Xu 等^[15]通过连续小波变换和灰度转换提取故障时频图,作为卷积神经网络的输入数据,并将网络中提取的特征输入到深度森林中进行分类。Liao 等^[16]设计了一种 Q 注意力机制来理解二次神经元,并将其应用于卷积神经网络进行故障诊断。

上述的方法尽管有不错的故障诊断效果,但是,都是使用的单分支 CNN,忽略了故障特征的复杂性,无法适用于不同场景的故障数据集。其次,忽略了分类器的选择对故障诊断的重要性。

近年来,多分支 CNN 在不同领域中应用,并展现出优越性。列如,Yan 等^[17]将多分支 CNN 应用于生物与医学领域,并证明相比传统的 ML 算法,多分支 CNN 更有效性。Wang 等^[18]通过在 PCANet 上增加一个基于 NPE 的卷积分支,并应用于计算机视觉领域中的图像分类任务,其性能远优于 PCANet。此外在故障诊断领域中,Huang 等^[19]在三分支一维卷积中加入 ECA 模块,Jia 等^[20]提出一种基于 Gramian 的新型降噪策略来增强振动信号中的特征信息,并且使用包含 RSB、TEB 和 FEB 三个分支的卷积进行轴承故障诊断。

同时,在众多分类器中,深度森林(Deep Forest, gcForest)^[21]作为一种新兴的集成学习方法,在不同领域上获得广泛的关注。它是由多粒度扫描与级联森林结合而成。由于级联森林通过集成多层决策树来进行特征学习和分类,且不依赖反向传播,这使得它在处理小样本数据集上展现出比传统方法更强的适应性和性能。但是像 Xu 等^[15]直接将 CNN 与深

度森林简单结合得到的故障诊断结果并非最佳。因为深度森林中多粒度扫描需要为不同的数据集重新设计和调整参数,并且直接使用 CNN 与深度森林结合会增加模型的复杂度。

因此,本文提出了多分支 CNN 与改进级联森林相结合的混合模型,命名为 CNN-ImpCasForest。首先将原始数据处理成二维的图像;其次,设计了一个多分支 CNN,并将二维图像输入多分支 CNN 中提取特征,通过多样化的特征学习和特征融合,提升了对复杂数据的特征提取能力;然后将提取的特征数据输入级联森林中,并将级联森林中的一个随机森林替换为极端梯度提升。这种部分替换结合了不同算法的优势,优化了模型整体性能,有效的提升了故障诊断的正确率。

这种混合模型不仅结合了深度学习和集成学习的优势,同时优化了深度森林的结构,去掉了多粒度扫描,降低了混合模型的复杂度。本文在不同场景的多个数据集进行了验证,实验表明,本文方法能有效提高故障诊断的准确率,并能更好的适用于不同场景的数据集。

1 资料与方法

1.1 数据预处理

数据预处理主要功能是从原始一维振动数据中提取特征,这些特征对故障诊断结果有很大影响。为了不增加数据预处理的复杂度,我们选择最常用的信号图像转换方法^[10]。该数据预处理方法无需预设参数,能够尽可能减少对专家经验的依赖,同时保留了原始信息的二维特征。因为本文网络使用 2×2 的滤波器,每层图像特征的尺寸都会减半,因此,本文选择 32×32 的尺寸。一维信号转换为二维图像的过程如图 1 所示。

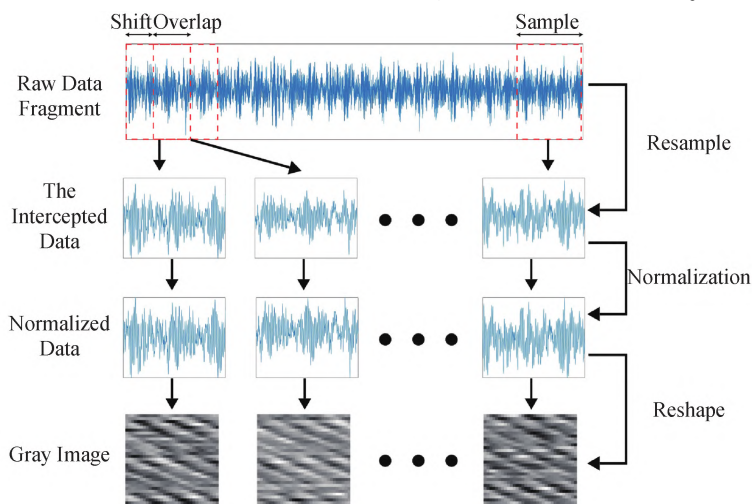


图 1 数据预处理步骤

如图 1 所示,为了获取 32×32 大小的图像,从原始信号中截取长度为 1 024 长度的片段信号,截取的间隔为 I 。令 $L(i)$ 表示信号片段的值,其中 $i = 1, 2, \dots, 1\,024$ 。 $P(j, k)$ 表示像素 (j, k) 处的强度值,其中 $j = 1, \dots, 32, k = 1, \dots, 32$, 公式如下

$$P(j, k) = \text{round} \left\{ \frac{L((j-1) \times 32 + k) - \text{Min}(L)}{\text{Max}(L) - \text{Min}(L)} \times 255 \right\}. \quad (1)$$

其中,函数 $\text{round}()$ 是取整函数, $\text{Max}(L)$ 是取信号片段 L 中的最大值, $\text{Min}(L)$ 是取信号片段 L 中的最小值,并且所有像素值都归一化到了 $[0, 255]$ 。

1.2 本文所提出的故障诊断模型

1.2.1 多分支卷积神经网络

为了提高模型的鲁棒性和性能,适应复杂任务需求,相对于单分支卷积神经网络能获得更好的故障特征,本文设计了一个多分支 CNN 网络。该网络

采用双分支设计,在保持轻量级网络特性的同时,提升特征提取能力和计算效率。2 个分支使用不同的卷积核大小,使模型能够关注局部和全局特征,从而捕获多尺度信息,提高复杂模式的识别率。并且这种结构有利于信息融合,通过结合不同来源的特征来提高分类准确率,降低过拟合的风险。与其他的多分支模型相比,双分支设计降低了模型复杂度和训练参数数量,提高了计算效率,确保了更稳定和收敛的训练过程。本文设计的多分支 CNN 网络结构如图 2 所示。

由图可知,该网络由 4 个卷积层 (C1、C2、C3 和 C4)、4 个池化层 (S1、S2、S3 和 S4)、2 个批量归一化层、3 个 dropout 层、2 个全局平均池化层 (G1 和 G2)、1 个连接层 G3、1 个全连接层 D1 和 1 个用于分类的 softmax 分类器组成。下面对使用的 dropout 层、全局平均池化层、激活函数和批量归一化层逐一介绍,其他卷积结构可以参考文献 [22]。

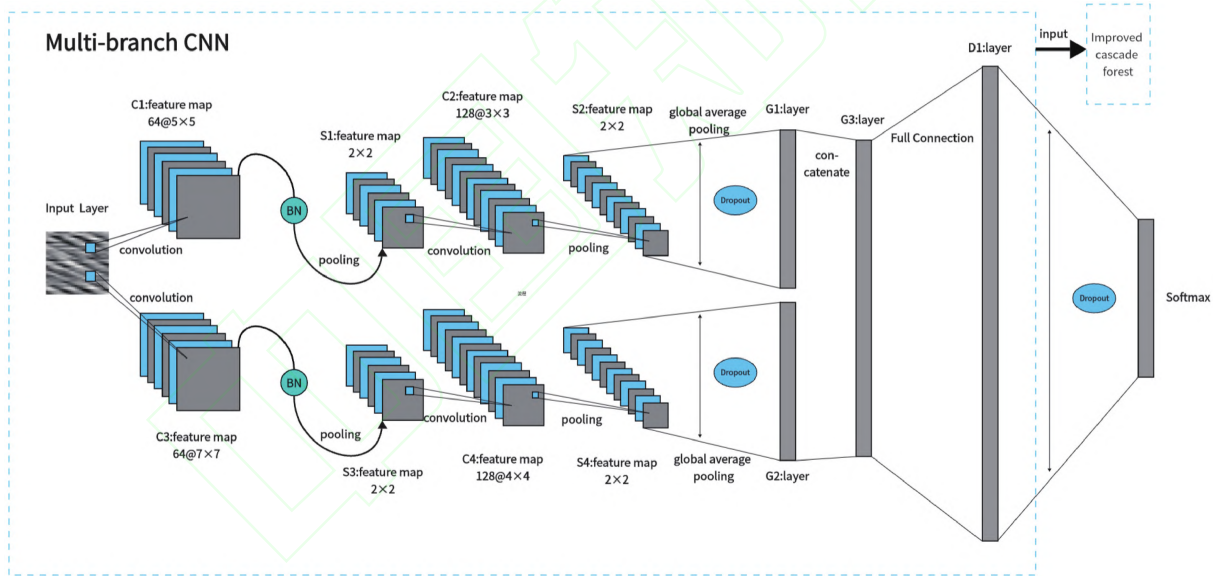


图 2 提出的多分支卷积神经网络结构

1.2.1.1 dropout 层

在模型中,dropout 函数^[23]设置为 0.3,以减少神经网络过拟合的倾向。在应用 dropout 层时,前向传播公式变为

$$r_j^{(l)} \sim \text{Bernoulli}(p), \quad (2)$$

$$\tilde{y}^{(l)} = r^{(l)} * y^{(l)}, \quad (3)$$

$$z_i^{(l+1)} = w_i^{(l+1)} \tilde{y}^{(l)} + b_i^{(l+1)}, \quad (4)$$

$$y_i^{(l+1)} = f(z_i^{(l+1)}). \quad (5)$$

这里, $r_j^{(l)}$ 是概率向量; $\text{Bernoulli}(p)$ 生成一个随机的 0 和 1 的向量; $\tilde{y}^{(l)}$ 表示修改后的激活函数; $y^{(l)}$ 是初始激活函数; $z_i^{(l+1)}$ 是卷积神经网络中下一

层的权重。

1.2.1.2 全局平均池化层

在卷积神经网络中,我们利用全局平均池化层^[24]来代替展平层,以便更好地保留空间信息,降低过度拟合的风险,尤其是对于单通道灰度图像。

1.2.1.3 激活函数

为了更灵活地处理数据并适应不同的数据分布和特征变化,从而提高模型的稳定性和泛化能力,所有卷积神经网络层都采用了 PReLU 函数^[25]作为激活函数。PReLU 的公式如下

$$f(y_i) = \begin{cases} y_i, y_i > 0, \\ a_i y_i, y_i \leq 0. \end{cases} \quad (6)$$

其中,负输入部分的斜率(a_i)是一个可学习的参数, y_i 是第*i*个通道上非线性激活函数*f*的输入。

1.2.1.4 批归一化层

批归一化(BN)^[26]是一种旨在提高深度学习模型训练速度和稳定性的技术。首先,计算均值

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i$$

和方差

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2,$$

然后将输入标准化 $\tilde{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}}$, 其中 ε 是常数。

最后通过可学习参数 γ 和 β 进行缩放和平移,得到 $y_i = \gamma \tilde{x}_i + \beta$ 。

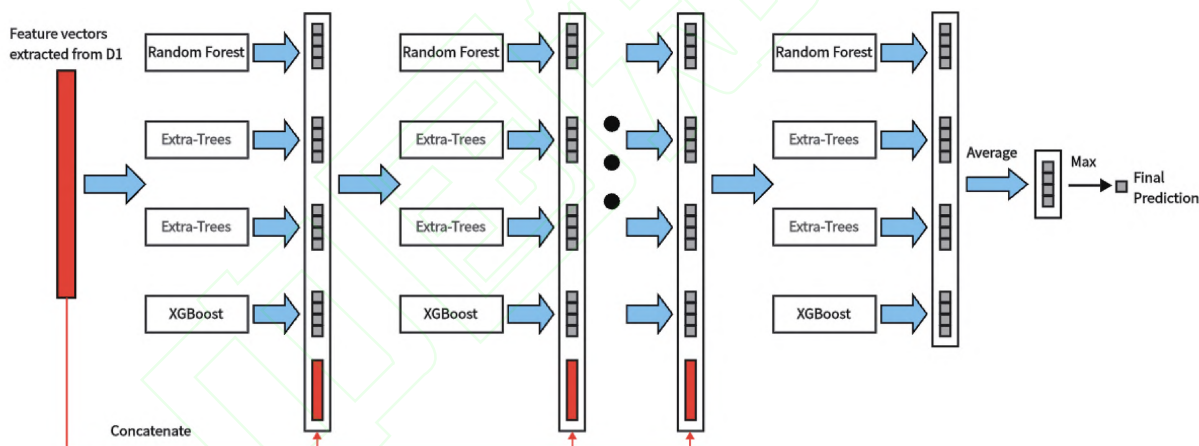


图 3 改进的级联森林结构

极端梯度提升^[27]是一种高效灵活的提升算法,是梯度提升决策树(GBDT)的全面升级。它通过特征分割构建决策树,能够识别特征空间内的非线性边界并有效捕获数据中的非线性关系。极端梯度提升还结合了正则化技术来控制模型复杂性并防止过度拟合。由于其在处理复杂数据集方面的有效性,它被广泛应用于故障诊断。

2 结果与分析

本研究利用 3 个轴承数据集和 1 个转子数据集来评估所提模型的性能。准确率用于评估模型性能,公式如下

$$ACC = \frac{TP}{N}. \quad (7)$$

1.2.2 改进的级联森林

级联森林^[21]是一种基于决策树的集成学习方法,旨在提升传统森林模型的性能。级联森林的设计思路源于深度学习的层次结构,但它是通过级联的方式将多个随机森林或极端随机森林层层堆叠,每一层都会基于上一层的输出特征进行训练。并且级联森林作为分类器具有较高的模型性能、简单的结构、较少的超参数调节需求,以及对小样本数据集的良好适应性。

改进的级联森林的结构如图 3 所示,它使用全连接层 D1 提取的特征作为输入数据,逐层进行训练。每层由 1 个随机森林,2 个极端随机森林,以及替代随机森林的极端梯度提升的学习算法的集合组成,这些算法经过训练后可获得新的特征向量。然后将这些获得的特征向量与输入数据连接起来,输入到下一层。此外,每层都采用 *k* 折交叉验证。如果连续训练 3 层后验证集上的性能没有提高,则停止训练。

其中 *TP* 表示正确分类的样本数, *N* 代表样本总数, *ACC* 表示准确率。

实验使用 Python 3.9 版、TensorFlow 2.6.0 版、NumPy 1.23.4 版和 Keras 2.6.0 版进行。计算机配置包括 AMD 锐龙 R7-5800H(配备 Radeon Graphics 3.20 GHz CPU)、16.0 GB RAM、1 TB SSD 和 Windows 11 操作系统。

2.1 数据集

实验进取的数据集如表 1 所示。第一个数据集是凯斯西储大学(CWRU)轴承数据集^[28]。所用的实验数据是在负载为 0 马力(HP)和转速为 1 797 rpm 时从驱动端的传感器收集的。数据集包含各种故障类型的样本,包括正常数据、内圈故障、外圈故障和滚动体故障,故障尺寸分别为 0.177 8、0.355 6、0.533 4 mm。在

数据预处理过程中,设置了 400 的采样间隔,并向数据集添加了不同强度的高斯白噪声。

第二个数据集为江南大学轴承数据集^[29],数据采样率为 50 kHz,所用实验数据是在 600 rpm 和 800 rpm 的运行速度下采集的,轴承状态包括内圈故障、外圈故障、滚珠故障、正常状态等 4 种,数据预处理时设置采样间隔为 1 024。

第三个数据集是渥太华大学的轴承数据集^[30]。数据是在转速从 15.2 Hz 增加到 26.7 Hz 时收集的。健康状态有 3 种:正常、内圈缺陷和外圈缺陷。在数据预处理期间,设置了 1 024 的采样

间隔。

第四个数据集来自武汉大学,包含转子数据集^[31]。实验数据采集速度为 1 200 rpm,采样频率为 2 048 Hz。数据集包括 4 种故障类型:正常、摩擦、不平衡和错位。在数据预处理期间,设置了 1 024 的采样间隔。

第五个数据集选用了江南大学轴承数据集和武汉大学转子数据集进行混合。数据集包括 8 种故障类型:轴承的内圈故障、正常状态、外圈故障和滚珠故障以及转子的正常、摩擦、不平衡和错位。在数据的预处理期间,设置了 1024 的采样间隔。

表 1 数据集

数据集	名字	类别数	每类的样本数	样本大小
凯斯西储大学轴承数据集	数据集 C	10	300	32×32/1 024
江南大学轴承数据集	数据集 J1,J2	4	200	32×32/1 024
渥太华大学轴承数据集	数据集 O	3	300	32×32/1 024
武汉大学转子数据集	数据集 W	4	90	32×32/1 024
混合数据集	数据集 H	8	90	32×32/1 024

2.2 参数设置

在所有数据集上,本文模型对比的模型包括 LeNet-5、ADCNN^[11]、MA1DCNN^[12]、DFCNN^[13]、DRSN-CS^[14]、DRSN-CW^[14]、geForest^[15]、CNN-geForest^[15]、QCNN^[16]、CA-MCNN^[19]、MIXCNN2^[32]和 GT-FE-Net^[20]。此外,所有网络的 epochs 均设置为 50, batch_size 设置为 128。

实验表中参数 α 表示训练样本与验证样本之和占样本总数的比例,训练样本与验证样本比例为 8:2。

本文设计的多分支 CNN 如表 2 所示,单分支网络 Branch1 和 Branch2 分别参考表 3 和表 4。

表 2 多分支卷积网络模型参数

网络层	核大小	参数	网络层输出
输入层	—	—	32×32
卷积层 C1	64@ 5×5	51 840	64@ 28×28
卷积层 C3	64@ 7×7	46 464	64@ 26×26
批归一化层 B1	—	256	64@ 28×28
批归一化层 B2	—	256	64@ 26×26
最大池化层 S1	2×2	0	64@ 14×14
最大池化层 S3	2×2	0	64@ 13×13
卷积层 C2	128@ 3×3	92 288	128@ 12×12
卷积层 C4	128@ 4×4	144 000	128@ 10×10
最大池化层 S2	2×2	0	128@ 6×6
最大池化层 S4	2×2	0	128@ 5×5
dropout 层 DR1	—	0	128@ 6×6
dropout 层 DR2	—	0	128@ 5×5
全球平均池化层 G1	—	0	128
全球平均池化层 G2	—	0	128
连接层 G3	—	0	256
全连接层 D1	—	132 096	512
dropout 层 DR3	—	0	512

表 3 Branch 1 模型参数

网络层	核大小	参数	网络层输出
输入层	—	—	32×32
卷积层 C1	64@ 5×5	51 840	64@ 28×28
批归一化层 B1	—	256	64@ 28×28
最大池化层 S1	2×2	0	64@ 14×14
卷积层 C2	128@ 3×3	92 288	128@ 12×12
最大池化层 S2	2×2	0	128@ 6×6
dropout 层 DR1	—	0	128@ 6×6
全球平均池化层 G1	—	0	128
全连接层 D1	—	66 560	512
dropout 层 DR3	—	0	512

表 4 Branch 2 模型参数

网络层	核大小	参数	网络层输出
输入层	—	—	32×32
卷积层 C1	64@ 7×7	46 464	64@ 26×26
批归一化层 B1	—	256	64@ 26×26
最大池化层 S1	2×2	0	64@ 13×13
卷积层 C2	128@ 4×4	144 000	128@ 10×10
最大池化层 S2	2×2	0	128@ 5×5
dropout 层 DR1	—	0	128@ 5×5
全球平均池化层 G1	—	0	128
全连接层 D1	—	66 560	512
dropout 层 DR3	—	0	512

2.3 在数据集 C 上的实验

数据集 C 是故障诊断中最常用的数据集。它总共包含 10 个状态:正常(N)、3 个内圈故障直径(IR007、IR014、IR021)、3 个滚动体故障直径(B007、B014、B021)和 3 个外圈故障直径(OR007、

OR014、OR021)。为了评估模型在噪声数据上的性能,将-8 dB、-4 dB、0 dB、4 dB 和 8 dB 水平的高斯白噪声添加到数据集中。训练样本、验证样本和测试样本的比例为 0.56:0.14:0.30。

从表 5 可以看出,本文所提出的模型在不同噪声都优于其他模型。其准确率从-8 dB 到 8 dB 逐渐提高,在 4 dB 和 8 dB 噪声下准确率分别达到 91.22% 和 97.78%。只有 MIXCNN2、GTFE-Net、DRSN-CS 和 DRSN-CW 在 8dB 时达到 97.11%、96.44%、94.78%和 94.00%的准确率,其余模型均低于 90%。

表 5 在数据集 C 上不同噪声的分类准确率 单位:%

方法	-8 dB	-4 dB	0 dB	4 dB	8 dB	无噪声
LeNet-5	22.44	26.89	39.22	56.11	65.44	71.78
ADCNN	23.89	32.56	19.56	25.78	40.56	67.44
MA1DCNN	20.56	11.00	38.11	65.11	42.00	82.67
DFCNN	10.00	22.22	27.44	51.56	46.44	81.67
DRSN-CS	43.89	52.67	77.00	88.44	94.78	99.33
DRSN-CW	41.67	53.33	69.89	79.67	94.00	98.89
gcForest	17.22	27.44	35.57	40.44	49.11	69.33
CNN-gcForest	28.33	45.56	59.22	70.78	84.11	92.88
QCNN	30.78	39.89	50.22	66.56	86.56	93.22
CA-MCNN	32.78	49.11	64.00	82.11	89.44	89.56
MIXCNN2	45.22	58.00	76.56	90.11	97.11	99.33
GTFE-Net	35.33	53.11	76.22	89.56	96.44	98.78
CNN-ImpCasForest	45.78	58.00	77.78	91.22	97.78	99.44

由于存在 10 个类别,混淆矩阵可能无法有效展示模型之间的差异。因此,这里使用彩图进行可视化。图 4 清楚地表明了 8 dB 噪声下分类准确率排名前 4 位的模型分类情况。与其他算法相比,本研究提出的算法在滚动体故障分类中的错误明显较少。

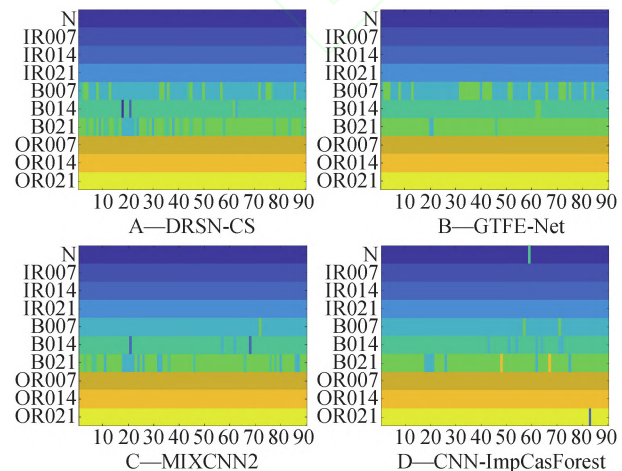


图 4 数据集 C 上 8 dB 噪声下的彩图

2.4 在数据集 J1 和 J2 上的实验

为了评估不同模型在不同工况下采集数据的性

能,本研究利用了以 600 rpm·min⁻¹和 800 rpm·min⁻¹采集的数据集 J1 和 J2。数据集包含 4 种状态:内圈故障 (IB)、正常 (N)、外圈故障 (OB) 和滚珠故障 (TB)。

从表 6 和表 7 可以看出,大部分模型在更高速度下采集的数据上实现了更高的准确率。此外,与其他算法相比,CNN-ImpCasForest 在不同的 α 值 (10%、30%、50% 和 70%)下的数据集 J1 上实现了最高的准确率。同样,在 J2 数据集上,CNN-ImpCasForest 在不同 α 值 (10%、30%、50% 和 70%)下均取得了最高的准确率,比次高准确率的算法分别高出 15%、3.93%、3.75% 和 0.83%。这表明 CNN-ImpCasForest 的表现明显优于其他算法,特别是在训练样本较低的情况下,其能适用于训练样本有限的场景。此外,与其他算法相比,它在以更高速度收集的数据集上获得了更好的准确率。

表 6 在数据集 J1 上不同 α 下的准确率 单位:%

方法	10%	30%	50%	70%
LeNet-5	44.58	48.21	54.50	52.08
ADCNN	37.22	48.75	38.50	48.33
MA1DCNN	24.03	51.96	27.00	43.33
DFCNN	40.83	31.96	36.00	37.92
DRSN-CS	48.19	72.14	73.25	82.92
DRSN-CW	46.39	71.07	74.50	82.50
gcForest	40.83	50.89	55.25	52.50
CNN-gcForest	52.22	58.04	64.25	64.58
QCNN	27.78	54.64	58.25	63.75
CA-MCNN	37.36	52.14	56.25	59.17
MIXCNN2	35.14	51.96	75.50	82.92
GTFE-Net	38.89	63.93	67.25	81.67
CNN-ImpCasForest	62.64	72.50	81.25	83.75

表 7 在数据集 J2 上不同 α 下的准确率 单位:%

方法	10%	30%	50%	70%
LeNet-5	54.86	59.64	55.00	55.42
ADCNN	51.81	29.82	42.50	50.83
MA1DCNN	25.00	42.50	71.75	68.33
DFCNN	25.14	34.82	27.25	29.58
DRSN-CS	62.08	82.68	88.75	90.42
DRSN-CW	55.56	81.07	87.75	90.83
gcForest	47.22	51.43	55.00	54.58
CNN-gcForest	51.39	76.61	80.00	91.25
QCNN	37.22	48.39	71.75	64.17
CA-MCNN	41.81	56.61	54.25	57.08
MIXCNN2	25.00	57.32	72.00	83.75
GTFE-Net	57.08	86.25	90.50	95.42
CNN-ImpCasForest	77.08	90.18	94.25	96.25

图5展示了在数据集J2上以 $\alpha = 70\%$ 实现较高准确率的4种算法的彩图。可以看出,CNN-ImpCasForest在识别外圈故障方面表现出色,与其他算法相比,内圈故障和滚珠故障的错误分类明显较少。

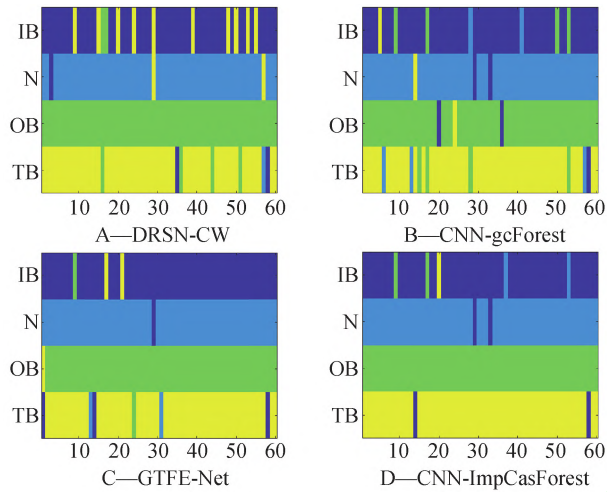


图5 数据集J2上 $\alpha = 70\%$ 下的彩图

2.5 在数据集O上的实验

数据集O是由加速条件下采集的数据组成,包含健康数据(HE)、内圈故障(IR)、外圈故障(OR)3种状态。

从表8可以看出,CNN-ImpCasForest在较低的 α 值下达到80.62%的准确率,比次高准确率的模型高出10%。这进一步体现了该模型适用于较小样本的训练。此外,它在其他训练比例下也达到了最高的准确率。

表8 在数据集O上不同 α 下的准确率 单位:%

方法	10%	30%	50%	70%
LeNet-5	61.85	66.35	63.78	71.48
ADCNN	51.36	56.19	59.56	57.41
MA1DCNN	51.73	60.79	66.67	75.56
DFCNN	33.58	54.28	55.56	54.44
DRSN-CS	33.46	76.83	80.44	86.30
DRSN-CW	50.40	71.90	81.11	87.04
gcForest	60.62	63.17	70.00	68.15
CNN-gcForest	70.62	78.10	84.89	83.70
QCNN	64.89	70.79	85.33	86.67
CA-MCNN	59.75	62.86	70.00	74.81
MIXCNN2	33.46	66.35	85.11	89.26
GTFE-Net	69.38	79.84	81.56	88.52
CNN-ImpCasForest	80.62	84.60	89.11	89.26

图6展示了在数据集O上,当 $\alpha = 30\%$ 时,准确率排名前4位的算法的彩图。可以看出,CNN-ImpCasForest在识别内圈故障其他故障方面则明显优

于其他算法。

2.6 在数据集W上的实验

转子是动力机械和工作机械中的关键部件,由轴承支撑旋转。为了验证各种模型在转子故障检测中的有效性,本文使用数据集W进行实验。该数据集包括4种类型的故障:正常、摩擦、不平衡和错位,分别缩写为NM、RU、UB和MA。

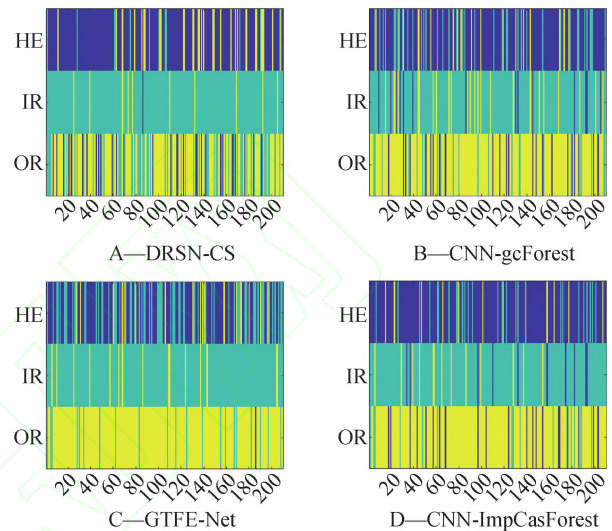


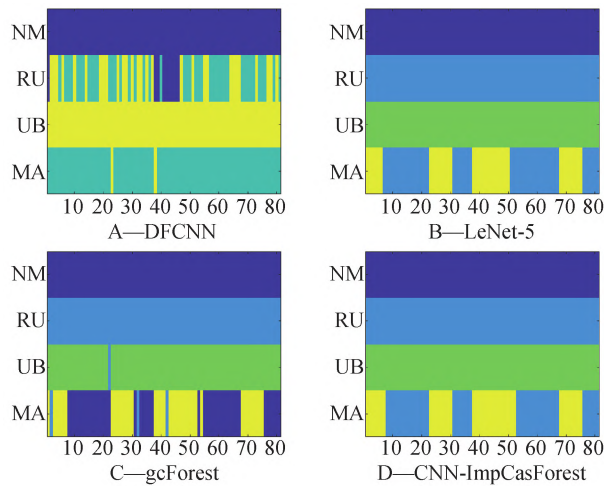
图6 数据集O上 $\alpha = 30\%$ 下的彩图

从表9可以看出,大部分算法都取得了不错的效果。CNN-ImpCasForest在 $\alpha = 10\%$ 时依然是所有算法中准确率最高的,在其他 α 取值时也达到了100%的准确率。说明本研究提出的算法能够很好地应用于转子数据集。

表9 在数据集W上不同 α 下的准确率 单位:%

方法	10%	30%	50%	70%
LeNet-5	85.80	98.02	100.00	100.00
ADCNN	48.46	82.94	75.00	93.52
MA1DCNN	25.00	25.00	49.44	54.63
DFCNN	63.27	50.00	99.44	100.00
DRSN-CS	25.00	23.02	100.00	100.00
DRSN-CW	25.00	25.00	100.00	100.00
gcForest	86.11	100.00	100.00	100.00
CNN-gcForest	74.07	100.00	100.00	100.00
QCNN	25.00	25.00	99.44	100.00
CA-MCNN	25.00	25.79	98.89	100.00
MIXCNN2	25.00	25.00	98.89	100.00
GTFE-Net	25.00	49.21	98.89	100.00
CNN-ImpCasForest	86.73	100.00	100.00	100.00

从图7可以看出,在 $\alpha = 10\%$ 的数据集W上,LeNet-5和CNN-ImpCasForest都只会错误识别错位类别中的故障。此外,CNN-ImpCasForest在识别错位故障方面表现出更高的准确率。

图 7 数据集 W 上 $\alpha=10\%$ 下的彩图

2.7 在数据集 H 上的实验

为了验证模型在多种故障类型上的有效性,本研究选用了江南大学数据集和武汉数据集进行混合,涵盖了轴承的内圈故障 (IB)、正常 (N)、外圈故障 (OB) 和滚珠故障 (TB) 以及转子的正常 (NM)、摩擦 (RU)、不平衡 (UB) 和错位 (MA)。

如表 10 和图 8 所示, CNN-ImpCasForest 在混合数据集上表现出比其他模型更好的效果。能够处理不同故障同时发生的情况,体现了其在实际工业应用中的潜力。

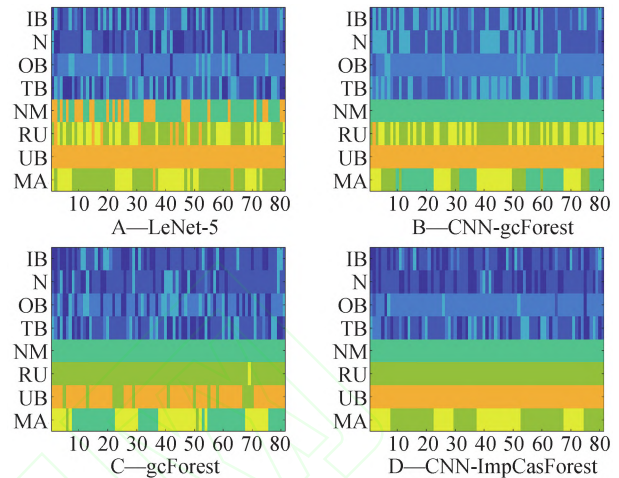
表 10 在数据集 H 上不同 α 下的准确率 单位: %

方法	10%	30%	50%	70%
LeNet-5	54.63	71.83	74.17	73.15
ADCNN	29.17	30.16	26.39	33.80
MA1DCNN	25.93	37.30	37.78	29.17
DFCNN	32.10	61.31	68.89	72.69
DRSN-CS	28.40	81.75	84.17	82.87
DRSN-CW	38.73	80.95	80.56	85.20
gcForest	61.88	73.61	72.78	69.91
CNN-gcForest	60.80	73.81	80.83	80.56
QCNN	25.93	62.90	50.83	76.85
CA-MCNN	34.26	53.37	72.78	77.31
MIXCNN2	25.77	73.02	73.33	81.02
GTFE-Net	36.88	81.94	84.17	77.31
CNN-ImpCasForest	67.90	81.94	84.44	85.65

2.8 消融实验

为了探究 CNN-ImpCasForest 中的多分支卷积神经网络是否能够有效提取更多特征,以及 ImpCasForest 作为分类器是否有效,本研究在数据集 C 上进行消融实验,采用 Branch1-ImpCasForest 和 Branch2-ImpCasForest 验证卷积神经网络的特征提

取能力,采用 CNN-Softmax、CNN-CasForest、CNN-XGBoost4、CNN-KNN 和 CNN-SVM 评估 ImpCasForest 作为分类器的有效性。

图 8 数据集 H 上 $\alpha=10\%$ 下的彩图

在本实验中,使用的分类器 CasForest 是由 2 个随机森林和 2 个完全随机森林组成。XGBoost4 用极端梯度提升替换了级联森林中的所有森林。KNN 是用 3 个最近邻居的 K 最近邻。SVM 是用线性核函数的支持向量机分类器。所有级联森林中每个森林有 101 棵树。

从表 11 可以看出,与 Branch1 和 Branch2 相比,所提出的多分支卷积神经网络的准确率提升了约 2%。为了更好地展示这 3 种卷积的特征提取能力,我们使用 t-SNE 可视化将提取的特征投影到三维空间中。如图 9 所示,本文方法可以明显地将同类数据聚类在一起,并将不同类型的数据分开,而 Branch1 和 Branch2 则无法有效区分某些数据点。此外,与仅使用随机森林和极端森林相比,使用 XGBoost 的级联森林可以将准确率提高 1% 到 2%。同时, CNN-ImpCasForest 在准确率方面明显优于 CNN-Softmax。

表 11 在数据集 C 上不同噪声下分类准确率 单位: %

方法	-8 dB	-4 dB	0 dB	4 dB	8 dB	无噪声
Branch1-ImpCasForest	44.78	55.56	75.11	89.44	96.44	98.33
Branch2-ImpCasForest	44.44	57.00	72.78	90.11	96.00	98.33
CNN-Softmax	29.78	25.89	54.78	39.00	82.78	94.67
CNN-XGBoost4	42.22	55.67	75.00	90.67	95.56	97.56
CNN-CasForest	44.78	56.78	75.44	91.11	96.78	98.11
CNN-KNN	41.56	55.33	74.11	90.89	96.67	97.78
CNN-SVM	43.67	55.89	76.33	90.78	96.89	98.78
CNN-ImpCasForest	45.78	58.00	77.78	91.22	97.78	99.44

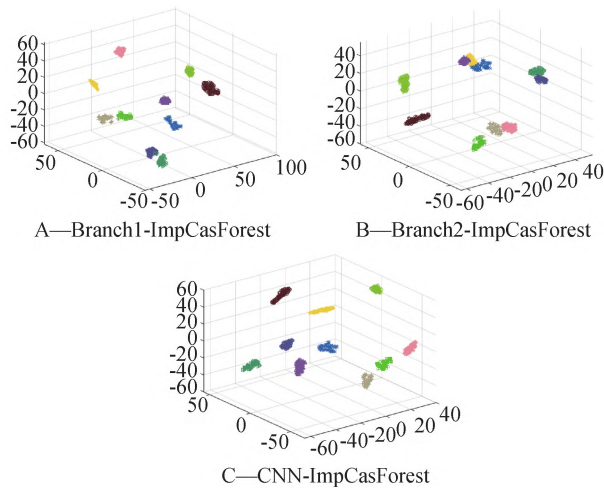


图9 在无噪声数据集 C 上可视化 D1 层输出的数据特征

3 结论

针对轴承和转子的故障诊断,其模型往往采用单分支 CNN,无法适用于不同场景的故障数据集,同时也忽略了分类器的选择对故障诊断的重要性。本文提出了一种结合多分支 CNN 与改进级联森林的混合故障诊断模型,名为 CNN-ImpCasForest。该模型以预处理后的灰度图像作为输入,通过多分支卷积神经网络提取特征信息,并将提取的特征直接输入一个将随机森林替换为极端梯度提升的级联森林中。这种部分替换结合了不同算法的优势,优化了模型整体性能,分类准确率比单纯使用极端梯度提升要好,同时删除了多粒度扫描,减少了模型的复杂度。

通过在不同状况下与其他模型对比,本文模型能够对噪声数据和小样本数据具有良好的适应性,并且在少量训练数据的情况下保持高效的学习能力,这充分体现本文模型具有较强的适应性和稳健性。同时,在所有数据集上获得最高诊断准确率,具有良好的泛化能力。

参考文献 (References)

- [1] ZHU Z Q, LEI Y B, QI G Q, et al. A review of the application of deep learning in intelligent fault diagnosis of rotating machinery [J]. *Measurement*, 2023, 206: 112346.
- [2] KIBRETE F, WOLDEMICHAEL D E, GEBREMEDHEN H S. Multi-Sensor data fusion in intelligent fault diagnosis of rotating machines: A comprehensive review [J]. *Measurement*, 2024: 114658.
- [3] LEI Y G, YANG B, JIANG X W, et al. Applications of machine learning to machine fault diagnosis: A review and roadmap [J]. *Mechanical Systems and Signal Processing*, 2020, 138: 106587.

- [4] GAWDE S, PATIL S, KUMAR S, et al. Multi-fault diagnosis of industrial rotating machines using data-driven approach: A review of two decades of research [J]. *Engineering Applications of Artificial Intelligence*, 2023, 123: 106139.
- [5] WANG T Z, QI J, XU H, et al. Fault diagnosis method based on FFT-RPCA-SVM for cascaded-multilevel inverter [J]. *ISA Transactions*, 2016, 60: 156–163.
- [6] GAO S Z, XU L T, ZHANG Y M, et al. Rolling bearing fault diagnosis based on SSA optimized self-adaptive DBN [J]. *ISA Transactions*, 2022, 128: 485–502.
- [7] CHEN J L, LI Z P, PAN J, et al. Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review [J]. *Mechanical Systems and Signal Processing*, 2016, 70: 1–35.
- [8] YIN Z Y, HOU J. Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes [J]. *Neurocomputing*, 2016, 174: 643–650.
- [9] PANDYA D H, UPADHYAY S H, HARSHA S P. Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN [J]. *Expert Systems with Applications*, 2013, 40 (10): 4137–4145.
- [10] 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*, 2017, 65(7): 5990–5998.
- [11] ISLAM M M M, KIM J M. Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network [J]. *Computers in Industry*, 2019, 106: 142–153.
- [12] WANG H, LIU Z L, PENG D D, et al. Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis [J]. *IEEE Transactions on Industrial Informatics*, 2019, 16 (9): 5735–5745.
- [13] ZHANG J Q, SUN Y, GUO L, et al. A new bearing fault diagnosis method based on modified convolutional neural networks [J]. *Chinese Journal of Aeronautics*, 2020, 33(2): 439–447.
- [14] ZHAO M H, ZHONG S S, FU X Y, et al. Deep residual shrinkage networks for fault diagnosis [J]. *IEEE Transactions on Industrial Informatics*, 2019, 16 (7): 4681–4690.
- [15] 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.
- [16] LIAO J X, DONG H C, SUN Z Q, et al. Attention-embedded quadratic network (qttnet) for effective and interpretable bearing fault diagnosis[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1–13.
- [17] YAN J L, ZHANG B, ZHOU M L, et al. Multi-Branch-CNN: Classification of ion channel interacting peptides using multi-branch convolutional neural network [J]. Computers in Biology and Medicine, 2022, 147: 105717.
- [18] WANG J P, RAN R S, FANG B. Global and local structure network for image classification[J]. IEEE Access, 2023, 11: 27963–27973.
- [19] HUANG Y J, LIAO A H, HU D Y, et al. Multi-scale convolutional network with channel attention mechanism for rolling bearing fault diagnosis [J]. Measurement, 2022, 203: 111935.
- [20] JIA L, CHOW T W S, YUAN Y. GTFE-Net: A gramian time frequency enhancement CNN for bearing fault diagnosis[J]. Engineering Applications of Artificial Intelligence, 2023, 119: 105794.
- [21] ZHOU Z H, FENG J. Deep forest[J]. National Science Review, 2019, 6(1): 74–86.
- [22] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84–90.
- [23] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929–1958.
- [24] LIN M, CHEN Q, YAN S C. Network in network[C]//Proceedings of the IEEE International Conference on Learning Representations, 2014.
- [25] HE K, ZHANG X Y, REN S Q, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]//Proceedings of the IEEE international conference on computer vision, 2015: 1026–1034.
- [26] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. arXiv preprint arXiv:1502.03167, 2015.
- [27] CHEN T Q, GUESTRIN C. Xgboost: A scalable tree boosting system[C]//Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016: 785–794.
- [28] LOPARO K. Case western reserve university bearing data center[J]. Bearings Vibration Data Sets, Case Western Reserve University, 2012: 22–28.
- [29] LI K, PING X L, WANG H Q, et al. Sequential fuzzy diagnosis method for motor roller bearing in variable operating conditions based on vibration analysis[J]. Sensors, 2013, 13(6): 8013–8041.
- [30] HUANG H, BADDOUR N. Bearing vibration data collected under time-varying rotational speed conditions [J]. Data in Brief, 2018, 21: 1745–1749.
- [31] LIU D, XIAO Z H, HU X, et al. Feature extraction of rotor fault based on EEMD and curve code[J]. Measurement, 2019, 135: 712–724.
- [32] ZHAO Z Q, JIAO Y H. A fault diagnosis method for rotating machinery based on CNN with mixed information [J]. IEEE Transactions on Industrial Informatics, 2022, 19(8): 9091–9101.

(责任编辑:郭芸婕)