

计算机集成制造系统
Computer Integrated Manufacturing Systems
ISSN 1006-5911, CN 11-5946/TP

《计算机集成制造系统》网络首发论文

题目：基于深度学习和多域决策融合的轴承故障智能诊断技术
作者：林诗麒，陈智丽，李宇鹏，孟维迎
收稿日期：2022-04-15
网络首发日期：2022-08-17
引用格式：林诗麒，陈智丽，李宇鹏，孟维迎. 基于深度学习和多域决策融合的轴承故障智能诊断技术[J/OL]. 计算机集成制造系统.
<https://kns.cnki.net/kcms/detail/11.5946.TP.20220816.1522.004.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于深度学习和多域决策融合的轴承故障智能诊断技术

林诗麒¹，陈智丽¹，李宇鹏²，孟维迎³

(1.沈阳建筑大学 信息与控制工程学院，辽宁 沈阳 110168；2.沈阳建筑大学 土木工程学院，辽宁 沈阳 110168；3.沈阳建筑大学 机械工程学院，辽宁 沈阳 110168)

摘要：滚动轴承是机械设备的关键部件，其振动信号的不平稳性及单一信息域特征的局限性，一定程度上增加了故障诊断难度。针对上述问题，提出一种基于深度学习和多域决策融合的轴承故障诊断技术。首先，应用 S 变换和递归图变换技术将振动信号从一维时域扩展至二维时频域和空间域；之后，为使诊断模型适应故障数据稀缺现状，构建泛化性和自适应性较好的微型卷积神经网络，学习提取信号的多域特征，网络参数低至 6 个数量级，可实现快速训练和故障诊断；最后，引入 D-S 证据理论对单域诊断结果进行融合。所提出的方法对凯斯西储大学数据集的 9 类轴承故障的平均诊断准确率达到 99.84%。

关键词：滚动轴承；微型卷积神经网络；多域融合；故障诊断

中图分类号：TP3

文献标识码：A

Intelligent bearing fault diagnosis technology based on deep learning and multi-domain decision fusion

LIN Shiqi¹, CHEN Zhili¹, LI Yupeng², MENG Weiyong³

(1. School of Information and Control Engineering, Shenyang Jianzhu University, Shenyang 110168, China; 2. School of Civil Engineering, Shenyang Jianzhu University, Shenyang 110168, China; 3. School of Mechanical Engineering, Shenyang Jianzhu University, Shenyang 110168, China)

Abstract: Rolling bearing is a key component of mechanical equipment. The instability of its vibration signal and the limitation of single domain features increase the difficulty of bearing fault diagnosis in some extent. On this basis, a bearing fault diagnosis technology based on deep learning and multi-domain decision fusion is proposed. Firstly, the S transform and recurrence plot transform are used to extend the vibration signal from one-dimensional time domain to two-dimensional time-frequency domain and spatial domain. Then, in order to adapt the diagnosis model to the lack of fault data, a micro-convolutional neural network with better generalization ability and adaptability is built to learn and extract multi-domain features of the signal, and the network parameters are as low as 6 orders of magnitude, which can be trained and classify fault data efficiently. Finally, D-S evidence theory is introduced to fuse the single domain diagnosis results. The proposed method achieves an average diagnostic accuracy of 99.84% for nine types of bearing faults in the Case Western Reserve University Dataset.

Keywords: rolling bearing; micro-convolutional neural network; multi-domain fusion; fault diagnosis

收稿日期：2022-04-15；修订日期：2022-07-18。Received 15 Apr. 2022; accepted 18 July 2022.

基金项目：工信部工业互联网创新发展工程——信息物理系统应用项目。**Foundation item:** Project supported by the Industrial Internet Innovation and Development Project of the Ministry of Industry and Information Technology -- Information Physics System Application Project, China.

1 引言

随着我国智能制造实力不断提升，现代化重工装置规模不断壮大，利用信息化技术对机械部件出现的故障进行精准诊断成为主要研究方向^[1]。滚动轴承作为关键旋转部件之一，决定着设备能否正常运作，对其故障实现快速智能诊断具有重大意义^[2]。

轴承故障诊断重点在于轴承振动信号的特征提取和分类识别，研究人员针对这两方面做了大量研究。在特征提取方面，从时间、频率、空间三个维度对轴承故障信号进行信息域变换。一维时域信号分解法^[3]分为有量纲和无量纲两类，有效值和峰值是常见的有量纲指标，分别描述轴承瞬时振动幅值的变化大小和异常突变值。无量纲指标，如裕度因子和峭度因子^[4]，对中高频和冲击性故障较为敏感，诊断低频故障信号表现较差。频域分析方法中的傅里叶变换^[5]主要作用于平稳信号，能得到较为详细的频域信息，但是缺失相关时域信息，应用受到限制。

针对单一时域或频域特征提取方法存在的不足，研究人员提出了一些基于二维空间的时频域分析方法，例如短时傅里叶变换^[6]、经验模态分解及衍生方法^[7-8]、小波变换及衍生方法^[9-10]等，当面向非线性、不平稳的复杂轴承振动信号时，时频域特征能够呈现更为丰富的故障信息且具备一定抗噪能力。从三维空间角度看，轴承故障信号是混沌系统的响应，可以映射到高维相空间，运动轨迹呈现出系统动力学特征。描述故障信号高维空间特征的方法有分形维数^[11]和近似熵^[12]等，这类方法可以表征系统整体概貌，有效度量轴承故障状态。

在分类识别方面，人工神经网络^[13]，支持向量机^[14]等传统机器学习方法对小数据集的分类诊断效果较好，但需建立复杂的人工特征提取工程；深度学习方法能够自适应提取轴承故障特征，更适用于多源、异构的工业数据。自编码器^[15]等主要用于无监督模式下的故障识别，卷积神经网络和循环神经网络^[16]主要用于有监督模式下的故障分类。文献[17]使用改进的卷积网络变体 AlexNet 网络诊断变工况轴承故障；文献[18]提出了一种全矢谱和循环神经网络的轴承故障诊断及寿命预测算法；文献[19]使用循环网络变体-双向长短时记忆网络对前后序列之间的关系建模，注重部分信号同整体信号的关联性以提高分类精度。

然而在工业生产中，轴承部件处于正常状态居多，采集和标注故障数据较为困难，解决这个问题通常从三方面入手，一是增广样本，例如文献[20]使用真实数据训练变分自编码器和生成对抗网络扩增样本数，再引入卷积神经网络进行诊断。二是迁移学习，例如文献[21]提出基于二维卷积和迁移学习的轴承故障诊断技术，将三种变速信号中的任意两种作为源域，余下一种作为目标域，诊断结果较为卓异。三是搭建结构小巧的深度学习网络模型，例如文献[22]使用仅有 11 层的

端到端卷积和长短时记忆网络，通过添加批标准层和舍弃层有效抑制过拟合，诊断小样本故障数据时行之有效。

综合上述研究，从二维和三维多维度出发，同时分析提取轴承振动信号的多域特征。鉴于故障信号的非线性和故障系统的混沌性，使用 S 变换(S Transform, ST)^[23]和递归图(Recurrence Plot, RP)变换^[24]分别对信号进行时频域和空间域变换。在二维空间中，S 变换的高斯窗函数能够自适应调节窗口大小，使短时傅里叶变换窗长固定的缺陷得以改善；在遇到小幅值高频信号时，S 变换具有较好的去噪效果，克服了经验模态分解易发生模态混叠效应的不足；此外，S 变换解决了小波变换方法中小波函数选择困难的问题。在三维空间中，RP 变换能处理长序列复杂信号，图形可视化空间拓扑结构，细微之处展现故障信号递归的空间相似性信息，和同领域的分形维数等方法相比，输入容量大，计算冗余度小。S 变换和 RP 变换虽将一维信号扩展至二维，但仅完成初级特征提取，深层特征学习和分类诊断任务由卷积神经网络实现。针对故障数据短缺，构建了具有自适应特征学习能力和较好泛化性的微型卷积神经网络(Micro-convolutional neural network, MCNN)，通过减少网络层数和引入全局平均池化等，使网络参数降至 6 个数量级，有效避免过拟合，使用 Softmax 函数得到信号的故障分类概率。考虑信号时频域和空间域特性的互补性，使用 D-S 证据理论融合两种信息域的诊断结果，减少结果的不确定性，在不影响诊断效率的同时提高诊断精度。

2 基于深度学习和多域决策融合诊断模型

本文提出的基于深度学习和多域决策融合诊断模型的工作流程如图 2.1 所示，在数据预处理阶段，使用最大最小归一化减小振动信号间的幅值差异，设置采样窗口为 512 划分样本；在多域变换阶段，分别对输入样本进行 S 变换和递归图变换，输出为二维时频域图像和空间域图像；在分类诊断阶段，为使样本适应神经网络输入，将图像变换到 256×256 大小并随机打散输入 MCNN 网络，使用五折交叉验证方法训练和评估模型，得到测试样本的初步诊断结果；在决策融合阶段，对得到的多域诊断结果使用 D-S 证据理论进行融合，减少结果的不确定性，从而提高诊断精度。

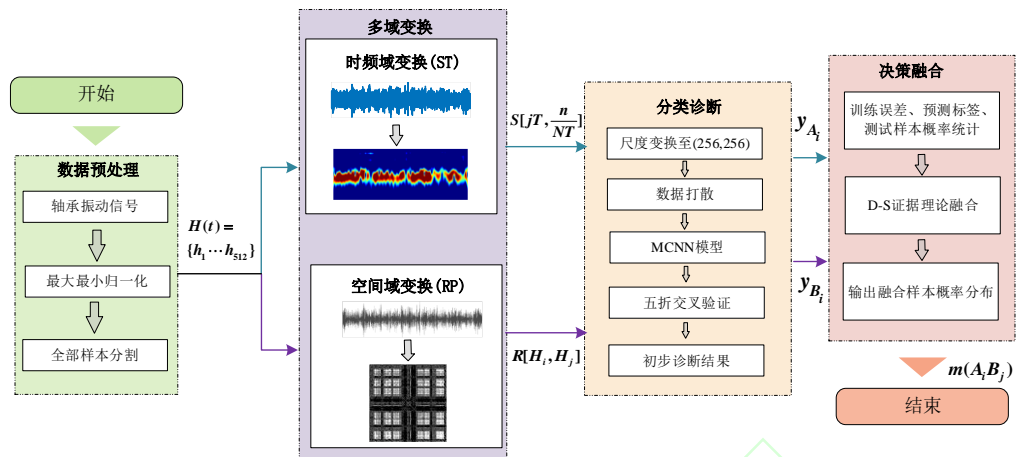


图 2.1 基于深度学习和多域决策融合的轴承故障诊断工作流程

2.1 轴承故障数据集

使用凯斯西储大学(Case Western Reserve University, CWRU)轴承数据中心提供的轴承故障振动信号作为实验数据，扭矩试验台(图 2.2)由 2 马力的 Reliance 电动机(左)、扭矩传感器(中)、规格为 6205-2RS JEM SKF 的深沟球轴承(中)、测力计(右)以及控制电子设备(未显示)组成。

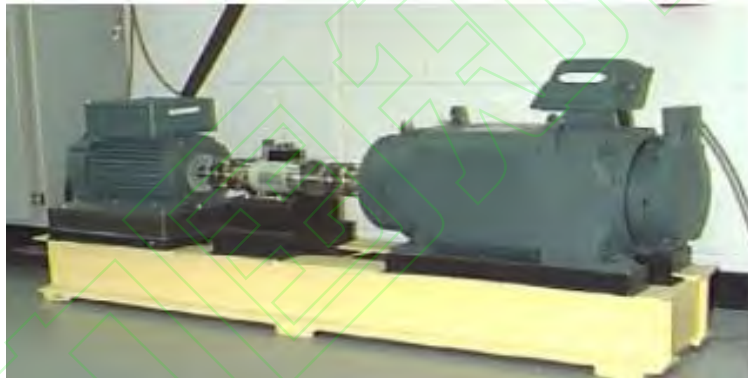


图 2.2 扭矩实验台

数据采集过程如下：首先使用电火花加工技术进行单点故障植入；然后再对每一类故障引入不同深度的断层；最后通过加速度传感器收集位于驱动端的故障数据，分别是：B07、IR07、OR07，B14、IR14、OR14，B21、IR21、OR21，代表着故障类型为滚动体(B)、内圈(IR)、外圈(OR)故障，损伤程度为 0.007 英寸、0.014 英寸、0.021 英寸的 9 类故障。本文使用当电机负载为 0HP、转速为 1797r/min、采样频率为 12kHz 时驱动端采集到的故障振动数据，选取 512 个数据点作为一个数据样本，共得到 2139 个样本，按照 3：1：1 划分训练集(1285 个)、验证集(427 个)、测试集(427 个)，确保样本集之间无交叉，细目如表 1。

表 1 实验样本集细目

故障类型	故障深度	总采样点个数	样本数
滚动体故障	0.007 英寸	122571	239
	0.014 英寸	121846	237
	0.021 英寸	121991	238
内圈点蚀故障	0.007 英寸	121265	236
	0.014 英寸	121846	237
	0.021 英寸	122136	238
外圈点蚀故障	0.007 英寸	122281	238
	0.014 英寸	121846	237
	0.021 英寸	122426	239

2.2 S 变换

S 变换使用高斯窗函数关联时域和频域信息并实现时间窗自适应调节，有效增强非线性信号的分辨率。设 T 为采样间隔， N 为采样长度， i 为虚数单位； m, n, j 分别为采样信号的时间编号、频率编号及频点，其取值范围是 $[0, N-1]$ 中的整数， $t = \frac{m}{NT}$ ，则数字信号 $H(t)$ 的离散 S 变换为：

$$S[jT, \frac{n}{NT}] = \begin{cases} \sum_{m=0}^{N-1} H(\frac{m+n}{NT}) e^{-\frac{2\pi^2 m^2}{n^2}} e^{-\frac{2\pi i m j}{N}}, n \neq 0 \\ \frac{1}{N} \sum_{m=0}^{N-1} H(\frac{m}{NT}), n = 0 \end{cases} \quad (1)$$

经公式(1)变换后，CWRU 九类故障部分样本的 S 变换时频图如图 2.3 所示，图像的长和宽分别代表时域和频域，大小为 $N \times \frac{N}{2}$ 。

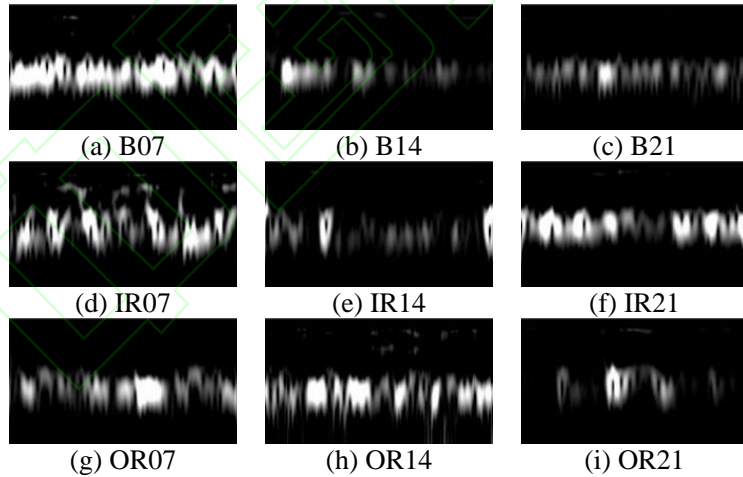


图 2.3 九类故障部分样本的 S 变换时频图

2.3 递归图变换

递归图变换通过延时嵌入将时序信号变换到相空间，再根据阈值 r 和相空间矢量间的欧氏距离 $d(H_i, H_j)$ 绘制由 0 和 1 组成的二维递归图像。使用互信息法^[25]和伪邻域法^[26]分别求解最佳延时系数 τ 及最小嵌入维度 m ，若一维时序信号为 $H(t) = \{h_1, h_2 \cdots h_N\}$ ，则重构矩阵 H 为：

$$H = \begin{bmatrix} h_1, h_{1+\tau} & \dots & h_{1+(m-1)\times\tau} \\ h_2, h_{2+\tau} & \dots & h_{2+(m-1)\times\tau} \\ \vdots & & \vdots \\ h_t, h_{t+\tau} & \dots & h_{t+(m-1)\times\tau} \end{bmatrix} \quad (2)$$

递归图像可由以下公式绘制，若 $r < d(H_i, H_j)$ ，则 $\text{Heaviside}(r - d(H_i, H_j)) = 0$ ，在递归图像上呈现为白点，反之为黑点，得到的递归图像维度为 $[N - \tau \times (m-1), N - \tau \times (m-1)]$ 。

$$R[H_i, H_j] = \text{Heaviside}(r - d(H_i, H_j)) \quad (3)$$

当输入长度为 512 的轴承振动信号时，九类故障部分样本空间重构后得到的二维递归图如图 2.4 所示，横轴和纵轴均为时间。

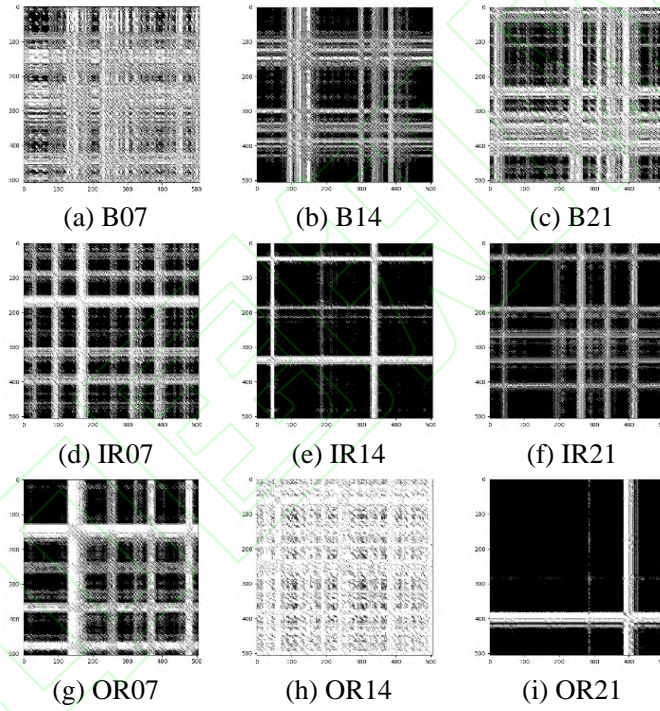


图 2.4 九类故障部分样本的递归图

2.4 微型卷积神经网络

深度学习模型往往具有庞大的网络结构，对高维度、大容量的数据有较强的处理能力，但是当输入样本数量较小时，复杂的深度学习模型通常得不到充分训练，会出现过拟合问题。本文所构建的微型卷积神经网络面向小样本数据，基于 VGG16 网络架构，继承深度学习自适应提取特征的优点，设计主要包含卷积层、最大池化层以及平均池化层的 9 层精简网络结构，缓解复杂学习模型参数冗余的问题，将网络训练参数降至 6 个数量级，旨在对网络每一层使用尽可能少的参

数，且在训练集和测试集上均取得不错的效果，提升模型的泛化性能，避免过拟合。表 2 给出了网络每一层属性设置、参数数量和输出维度等信息。

表 2 MCNN 各层属性设置

层数	结构类型	属性	参数数量	输出维度
1	输入(Input)	Size=256×256×1	/	/
2	第一卷积层(Conv_1)	Stride=1, filter=3×3×64, activation='ReLU'	640	254 × 254 × 64
3	第一最大池化层(Pool_1)	filter=2×2	0	127 × 127 × 64
4	第二卷积层(Conv_2)	Stride=1, filter=3×3×128, activation='ReLU'	73856	125 × 125 × 128
5	第二最大池化层(Pool_2)	filter=2×2	0	62 × 62 × 128
6	第三卷积层(Conv_3)	Stride=1, filter=3×3×512, activation='ReLU'	590336	60 × 60 × 512
7	第三最大池化层(Pool_3)	filter=2×2	0	30 × 30 × 512
8	全局最大平均池化层(GAP)	/	0	1 × 512
9	分类层(Softmax)	Neurons=9	4617	9

注：Stride 为步长；filter 为滤波器；Neurons 为神经元

如表 2 所示，受 VGG16 启发，在 MCNN 网络第二层添加由 64 个核函数组成的卷积层。核函数大小为 3×3，以多个小卷积核级联的方式代替大卷积核，扩大输出特征图映射域，减小计算量。选用线性整流函数(Rectified Linear Unit, ReLU)选择性激活输出单元，解决“梯度消失”问题。为了使输入与输出局部特征之间具有平移不变性，在卷积层后一层添加池化层，池化方式选择具有降维功能的最大池化。MCNN 的第 4-7 层，是在上一层基础上叠加了 2 组卷积—池化块。单层卷积捕获低阶特征，多层卷积整合低阶特征，获得高级特征，且层与层之间可实现权值共享^[27]。经过最后一次池化，特征图已经由原来的 254×254×64 变为 30×30×512。尽管此时得到的特征表示高阶抽象特征，但无法与输出建立联系，故添加全局平均池化层，通过加权、模板匹配将特征映射到各个类别。另外，由于该层结构对每个通道的每张特征图求平均，全程无需优化参数，相当于生成一个正则化器，网络泛化能力也会得到提高。经过上述操作后，网络得到各个故障类别的评价分数，为方便后期评估模型性能，在 MCNN 模型最后一层添加 Softmax 层，将范围是 $[-\infty, +\infty]$ 的分数归一化到[0,1]之间的概率值。

MCNN 模型配置主要涉及到计算损失、选择优化器。由于 MCNN 模型分类层是 Softmax 函数，使用交叉熵损失函数可以直接处理 Softmax 输出概率值，减少计算量。优化器是引导神经网络完成参数更新的工具，考虑到计算效率以及内存需求，使用自适应矩估计优化器，通过修正一阶、二阶动量偏差，自动调节学习率，提高模型收敛速度。

2.5 D-S 证据理论融合

D-S 证据理论^[28]能够处理由于传感器异质和噪声等带来的不确定性和模糊性信息，假设 A、B 是属于同一识别框架 Ω 下两个相互独立的传感器，则计算基本信任分配函数(Basic Belief

Assignment, BBA)得到新证据 m_1A 和 m_2B 的公式如下:

$$\begin{aligned} m_1(A_i) &= \frac{y(A_i)}{\sum_{i=1}^n y(A_i) + E_A} \\ m_2(B_j) &= \frac{y(B_j)}{\sum_{j=1}^n y(B_j) + E_B} \end{aligned} \quad (4)$$

n 代表类别数 9, E_A 、 E_B 代表分类器训练最小误差值。对于同一样本, 经过 MCNN 分类后, 会分别得到基于时频域和空间域的统计概率分布 $y(A_i)$ 和 $y(B_j)$, 统计概率分布中最大概率值对应的标签即为预测标签, 如果两个信息域的预测标签均等于真实标签, 则无需融合, 如果存在一个信息域的诊断结果有误, 则需按照合取规则融合有差异的信息:

$$m(A_i B_j) = \frac{\sum_{A_i \cap B_j = C} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)} \quad (5)$$

不确定性可以评估融合结果好坏, 设 $m(x_i)$ 为新的 BBA, 不确定性 $m(\theta)$ 的数学描述如下:

$$m(\theta) = 1 - \sum_{i=1}^n m(x_i) \quad (6)$$

3 实验与结果分析

在模型的测试过程中, 选用准确率(Accuracy, Acc)的平均值和标准差作为评价指标, 由式(7)和(8)计算, 其中, *True positive* 代表正确分类样本总数, *False negative* 代表错误分类样本总数; *std* 是标准值, n 为交叉验证折数, 这里等于 5; a_i 是第 i 折准确率, \bar{a} 是 n 折准确率平均值。

$$Acc = \frac{True\ positive}{True\ positive + False\ negative} \times 100\% \quad (7)$$

$$std = \frac{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2}}{\sqrt{n}} \quad (8)$$

设定 MCNN 网络的初始学习率为 0.001, 批处理大小为 32, 训练 100 轮次并进行五折交叉验证, 用验证集最佳权重在测试集上得到的基于时频域的 ST-MCNN 和基于空间域的 RP-MCNN 分类结果如图 3.1 和图 3.2 所示。

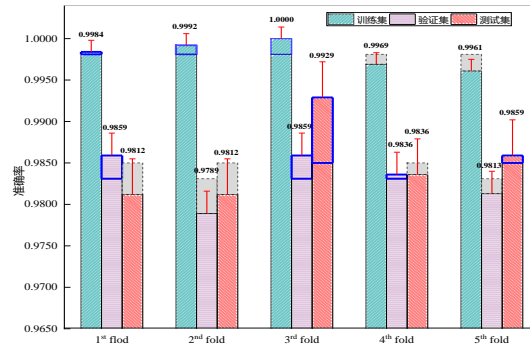


图 3.1 ST-MCNN 五折交叉验证结果

图中左斜线、横线及右斜线柱状图分别代表了训练集、验证集及测试集准确率，顶部实线和灰色虚线框分别为准确率高于和低于平均准确率部分，误差棒的长短象征着标准值的大小，误差棒越短，诊断模型的稳定程度越高。

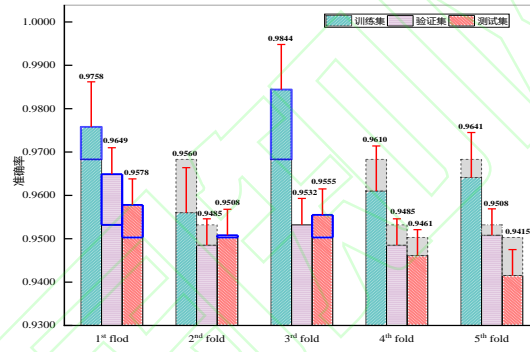


图 3.2 RP-MCNN 五折交叉验证结果

只用 MCNN 网络对未经过 ST 或 RP 变换的原始轴承故障信号诊断，在测试集上，相较于 ST-MCNN 平均准确率下降 9.95 个百分点；相较于 RP-MCNN 平均准确率下降 6.75 个百分点，结果如图 3.3 所示。

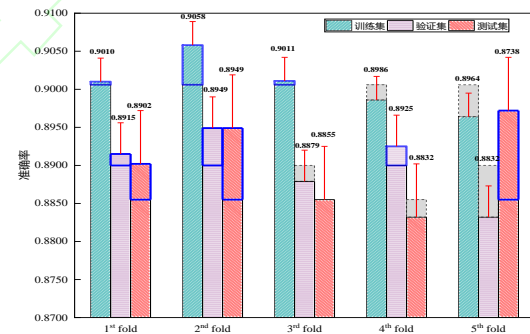


图 3.3 MCNN 五折交叉验证结果

对 ST-MCNN 和 RP-MCNN 网络诊断结果进行 D-S 证据理论融合，部分测试样本的原始预测概率、BBA 值、以及融合结果分别如表 3~7 所示。

表 3 ST-MCNN 部分测试样本原始概率

诊断模型	真实	预测	m(1)	m(2)	m(3)	m(4)	m(5)	m(6)	m(7)	m(8)	m(9)
ST-MCNN	3	3✓	0.0000	0.0000	0.9998	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000
	2	2✓	0.0000	0.9999	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	3✓	0.0000	0.0683	0.9315	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000
	9	9✓	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
	7	3✖	0.1822	0.0146	0.7216	0.0000	0.0000	0.0000	0.0816	0.0000	0.0000
	2	3✖	0.0000	0.4256	0.5744	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	1✖	0.3399	0.1995	0.2228	0.0017	0.0000	0.0003	0.2354	0.0004	0.0000
	1	1✓	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	9	5✖	0.0000	0.0013	0.0000	0.0000	0.6770	0.0000	0.0000	0.0000	0.3217
	3	1✖	0.6820	0.0013	0.1423	0.0002	0.0000	0.0000	0.1735	0.0007	0.0000

表 4 RP-MCNN 部分测试样本原始概率

诊断模型	真实	预测	m(1)	m(2)	m(3)	m(4)	m(5)	m(6)	m(7)	m(8)	m(9)
RP-MCNN	3	1✖	0.8645	0.0006	0.1344	0.0000	0.0000	0.0000	0.0005	0.0000	0.0000
	2	3✖	0.0067	0.0692	0.9241	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	2✖	0.0782	0.7430	0.1788	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	9	5✖	0.0000	0.0009	0.0000	0.0006	0.9790	0.0000	0.0000	0.0000	0.0122
	7	7✓	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.9998	0.0000	0.0000
	2	2✓	0.0000	0.9974	0.0007	0.0000	0.0019	0.0000	0.0000	0.0000	0.0000
	2	2✓	0.0000	0.9316	0.0091	0.0000	0.0001	0.0592	0.0001	0.0000	0.0000
	1	3✖	0.3380	0.0038	0.6582	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	9	9✓	0.0000	0.0092	0.0001	0.0000	0.0001	0.0000	0.0002	0.0000	0.9905
	3	3✓	0.0239	0.4129	0.5620	0.0000	0.0000	0.0000	0.0012	0.0000	0.0000

表 5 ST-MCNN 部分测试样本的 BBA 值

诊断模型	真实	预测	m(1)	m(2)	m(3)	m(4)	m(5)	m(6)	m(7)	m(8)	m(9)	m ($\theta\theta$)
ST-MCNN	3	3✓	0.0000	0.0000	0.9984	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0015
	2	2✓	0.0000	0.9986	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0014
	3	3✓	0.0000	0.0682	0.9302	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0016
	9	9✓	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9986	0.0014
	7	3✖	0.1819	0.0146	0.7206	0.0000	0.0000	0.0000	0.0815	0.0000	0.0000	0.0014
	2	3✖	0.0000	0.4250	0.5736	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0014
	2	1✖	0.3395	0.1992	0.2225	0.0017	0.0000	0.0003	0.2350	0.0004	0.0000	0.0014
	1	1✓	0.9986	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0014
	9	5✖	0.0000	0.0013	0.0000	0.0000	0.6761	0.0000	0.0000	0.0000	0.3212	0.0014
	3	1✖	0.6811	0.0013	0.1421	0.0002	0.0000	0.0000	0.1732	0.0007	0.0000	0.0014

表 6 RP-MCNN 部分测试样本的 BBA 值

诊断模型	真实	预测	m(1)	m(2)	m(3)	m(4)	m(5)	m(6)	m(7)	m(8)	m(9)	m ($\theta\theta$)
RP-MCNN	3	1✖	0.8581	0.0006	0.1334	0.0000	0.0000	0.0000	0.0005	0.0000	0.0000	0.0074
	2	3✖	0.0066	0.0687	0.9173	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0074
	3	2✖	0.0776	0.7375	0.1775	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0074
	9	5✖	0.0000	0.0009	0.0000	0.0006	0.9790	0.0000	0.0000	0.0000	0.0121	0.0074
	7	7✓	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.9925	0.0000	0.0000	0.0073
	2	2✓	0.0000	0.9900	0.0007	0.0000	0.0019	0.0000	0.0000	0.0000	0.0000	0.0074
	2	2✓	0.0000	0.9248	0.0090	0.0000	0.0001	0.0588	0.0001	0.0000	0.0000	0.0072
	1	3✖	0.3355	0.0037	0.6534	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0074
	9	9✓	0.0000	0.0091	0.0001	0.0000	0.0001	0.0000	0.0002	0.0000	0.9832	0.0073
	3	3✓	0.0009	0.0465	0.9449	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	0.0074

表 7 部分测试样本融合结果

真实	融合	m(1)	m(2)	m(3)	m(4)	m(5)	m(6)	m(7)	m(8)	m(9)	m ($\theta\theta$)
3	3✓	0.0091	0.0000	0.9908	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
2	2✓	0.0001	0.9832	0.0166	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
3	3✓	0.0006	0.2317	0.7677	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	9✓	0.0000	0.0001	0.0000	0.0000	0.0657	0.0000	0.0000	0.0000	0.9337	0.0005
7	7✓	0.0148	0.0012	0.0602	0.0000	0.0000	0.0000	0.9237	0.0000	0.0000	0.0001
2	2✓	0.0000	0.9892	0.0108	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	2✓	0.0125	0.9594	0.0186	0.0001	0.0000	0.0005	0.0088	0.0000	0.0000	0.0001
1	1✓	0.9973	0.0000	0.0027	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	9✓	0.0000	0.0001	0.0000	0.0000	0.0154	0.0000	0.0000	0.0000	0.9844	0.0001
3	3✓	0.0393	0.0009	0.9507	0.0000	0.0000	0.0000	0.0089	0.0000	0.0000	0.0002

对比表 5、6 和表 7，融合前，ST-MCNN 和 RP-MCNN 各有 5 个样本诊断错误，不确定性较高，诊断错误样本较多；融合后，10 组测试样本全部诊断正确，不确定性接近于 0，误差变小。

图 3.4 和图 3.5 给出两类模型错误样本分布情况，表明两模型间具有互补性，适合融合诊断。

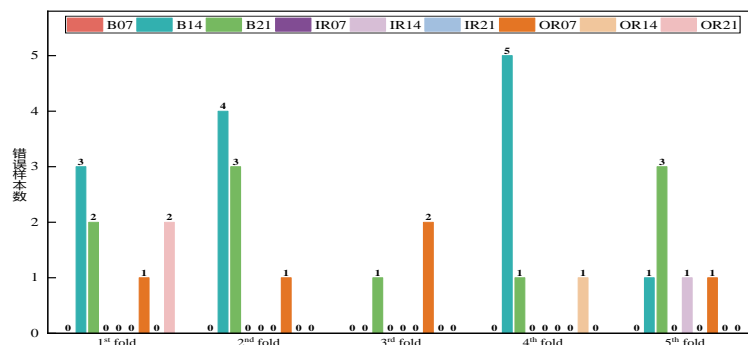


图 3.4 ST-MCNN 错误样本分布

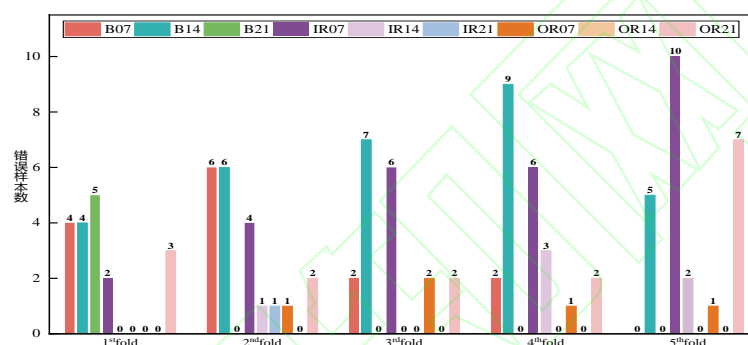


图 3.5 RP-MCNN 错误样本分布

从分类类别看，提取不同信息域特征会产生不同分类结果，对于类型为 B07、B14、IR07、IR14、IR21、OR21 故障，信号时频域特征对决策影响更大，ST-MCNN 诊断效果更佳(图 3.4)；对于 B21、OR14 故障，空间域信息较为强烈，使用 RP-MCNN 诊断时，错误分类样本数较少(图 3.5)。从测试结果看，在折数等于 1 时，ST-MCNN 模型共有 8 个样本诊断错误(图 3.4)，分别是第 130、163、176、204、246、274、392、399 段，其余段样本全部正确；RP-MCNN 模型有 18 个样本被误判(图 3.5)，分别为是第 10、11、59、72、101、106、117、171、200、238、290、307、333、359、378、384、420、423 段，两模型判断失误样本中无交叉。在第 2、3、4、5 折中，基于两信息域误判样本重叠数量分别为 0、1、1、2 个，诊断结果具有多样性。对两信息域的分类结果进行决策级融合，针对同一样本，将准确率较高模型的输出作为最终结果，实现互补。表 8 给出全部测试样本融合诊断准确率。

表 8 全部测试样本融合诊断准确率

折数	ST-MCNN 平均准确率	RP-MCNN 平均准确率	融合平均准确率
1 st fold	0.9812 (↑1.88%)	0.9578 (↑4.22%)	1.0000
2 nd fold	0.9812 (↑1.88%)	0.9508 (↑4.92%)	1.0000
3 rd fold	0.9929 (↑0.48%)	0.9555 (↑4.22%)	0.9977
4 th fold	0.9836 (↑1.41%)	0.9461 (↑5.16%)	0.9977
5 th fold	0.9859 (↑0.94%)	0.9415 (↑5.38%)	0.9953
Average±std	0.9850±0.0043	0.9503±0.0060	0.9984±0.0018

表 8 中，融合之后模型在准确率上相较于之前均有提高。较于 ST-MCNN，平均上升 1.32 个百分点；较于 RP-MCNN，成效更加显著，平均上升 4.78 个百分点。融合之后平均准确率达到 99.84%，标准值降至 0.0018。标准值离散程度减小，模型抗噪性能增强，诊断结果更加精确。为了验证本文所提出方法的有效性，选用其他五种方法对比。使用同一数据集，同一验证方法完成实验，表 9 是不同方法轴承故障诊断结果。

表 9 不同故障诊断方法对比

	方法一	方法二	方法三	方法四
文献来源	Althubiti 等 ^[29]	Khorram 等 ^[30]	Sadrezazami 等 ^[31]	Jiang 等 ^[32]
数据来源	CWRU	CWRU	CWRU	CWRU
诊断模型	VGG16	LSTM	时域数据+1D CNN	频域数据+SVM
五折交叉验证结果 ACC	1 st fold: 0.9672	1 st fold: 0.9479	1 st fold: 0.9414	1 st fold: 0.9113
	2 nd fold: 0.9765	2 nd fold: 0.9122	2 nd fold: 0.9180	2 nd fold: 0.9278
	3 rd fold: 0.9859	3 rd fold: 0.9395	3 rd fold: 0.9391	3 rd fold: 0.9018
	4 th fold: 0.9578	4 th fold: 0.9295	4 th fold: 0.9157	4 th fold: 0.8900
	5 th fold: 0.9578	5 th fold: 0.9372	5 th fold: 0.9297	5 th fold: 0.9013
	Average±std: 0.9690±0.0103	Average±std: 0.9313±0.0120	Average±std: 0.9288±0.0105	Average±std: 0.9064±0.0126

表 9 中 VGG16 模型的结构均较为复杂，训练参数总数高达 1500 万，对样本维度和数量要求较高，而实验样本较少，导致模型得不到充分训练，拟合能力变差，准确率降低。LSTM 模型通过“门控单元”的方式窥析信号在时间尺度上的相关性，选用双曲线正切函数作为激活函数，用 1 来初始化遗忘门偏差，添加 Dropout 层舍弃部分神经元，尽管增强了时域信号间的分析处理能力，但是在信号特征提取方面仍有所缺失，9 类故障的平均准确率为 93.13%。方法三重现文献[31]所述方法，使用自适应一维紧凑卷积神经网络(由三个卷积层，卷积方式：1D，卷积核尺度：1×9，卷积核个数：[60 40 40]；二个池化层，池化方式：最大池化，池化尺度：4，池化核个数：20；一个全连接层组成)分类，设定初始学习率为 0.001，选择随机梯度下降优化算法更新模型参数，对 CWRU 中 9 类故障进行诊断，经五折交叉验证后，平均准确率达到 92.88%。相较于本文，方法三未对时域信号做任何变换，故障特征不明显，诊断效果略显不足。方法四重现文献[32]工作，首先对故障信号做归一化预处理，然后进行 FFT 变换，得到单边频谱信息，最后构建 SVM 分类器，选用径向基核函数映射特征，惩罚因子为 3.2，核半径为 5.7。相较于本文使用深度学习方法进行分类，传统的机器学习算法 SVM 在很大程度上依赖于人工特征提取和特征选择，容错性较差，对 CWRU 中 9 类故障诊断平均准确率仅为 90.84%。

4 结束语

本文提出基于深度学习和信号时频域及空间域的决策融合诊断模型，分别用于诊断 9 类不同类别、不同损伤程度的轴承故障信号。实验结果表明，经 ST 变换和 RP 变换得到的二维时频域和空间域特征矩阵，蕴含着丰富的非线性故障信息；可训练数量低至 449669 个的 9 层 MCNN 网络层

次结构简洁, 功能完备, 能够提取故障信号有效特征, 对数据规模需求小, 适用于现阶段工业轴承故障数据异构、稀缺的实际情况; D-S 证据理论对两种诊断模型进行决策级融合, 减小由噪声等干扰带来的不确定性和模糊性, 实现诊断性能互补, 有效提高整体诊断精度和诊断结果可靠性。

参考文献:

- [1] ZHOU Ji—"Made in China 2025" main direction[J]. Chinese Mechanical Engineering, 2015,26(17): 2273-2284 (in Chinese). [周 济. 智能制造——“中国制造 2025”的主攻方向[J]. 中国机械工程, 2015,26(17): 2273-2284.]
- [2] PENG Cheng, HE Jing, TANG Zhaohui, et al. Early fault classification algorithm of rolling bearing based on two-dimensional EKNN[J]. Computer Integrated Manufacturing Systems, 2021,27(1): 12 (in Chinese). [彭 成,贺 婧,唐朝晖,等. 基于双维度 EKNN 的滚动轴承早期故障分类算法[J]. 计算机集成制造系统, 2021,27(1): 12.]
- [3] LI Wenfeng, DAI Haomin, XU Aiqiang. New time domain index and probabilistic neural network and their application in fault diagnosis of rolling bearing[J]. Mechanical Science and Technology for Aerospace Engineering, 2016, 35(9): 1382-1386 (in Chinese). [李文峰, 戴豪民, 许爱强. 时域新指标和 PNN 在滚动轴承故障诊断中的应用[J]. 机械科学与技术, 2016, 35(9): 1382-1386.]
- [4] Tian J, Morillo C, Azarian M H, et al. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis[J]. IEEE Transactions on Industrial Electronics, 2015, 63(3): 1793-1803.
- [5] Gao Y, Yu D, Wang H. Fault diagnosis of rolling bearings using weighted horizontal visibility graph and graph Fourier transform[J]. Measurement, 2020, 149: 107036.
- [6] LI Heng, ZHNAG Qin, SUN Yuantao. Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network[J]. Journal of Vibration and Shock, 2018, 37(19): 125-131 (in Chinese). [李 恒, 张 氢, 秦仙蓉, 孙远韬. 基于短时傅里叶变换和卷积神经网络的轴承故障诊断方法[J]. 振动与冲击, 2018, 37(19): 125-131.]
- [7] Zhang X, Liang Y, Zhou J. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM[J]. Measurement, 2015, 69: 164-179.
- [8] Zhang M, Jiang Z, Feng K. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump[J]. Mechanical Systems and Signal Processing, 2017, 93: 460-493.
- [9] Shen C, Wang D, Kong F, et al. Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier[J]. Measurement, 2013, 46(4): 1551-1564.
- [10] Zhang K, Ma C, Xu Y, et al. Feature extraction method based on adaptive and concise empirical wavelet transform and its applications in bearing fault diagnosis[J]. Measurement, 2021, 172: 108976.
- [11] Zhang Y, Ren G, Wu D, et al. Rolling bearing fault diagnosis utilizing variational mode decomposition based fractal dimension estimation method[J]. Measurement, 2021, 181: 109614.
- [12] WANG Yijing, TAN Haiyan. Research on fault diagnosis method of rolling bearing based on multi-dimension Cross Approximate Entropy[J]. Machine Design and Research, 2019,35(1):110-112+124 (in Chinese). [王宜静, 谭海燕. 基于多维度互近似熵的滚动轴承故障诊断方法研究[J]. 机械设计与研究, 2019, 35(1): 110-112+124.]

-
- [13] Bangalore P, Tjernberg L B. An artificial neural network approach for early fault detection of gearbox bearings[J]. IEEE Transactions on Smart Grid, 2015, 6(2): 980-987.
- [14] JIN Jiangtao, XU Zifei, LI Chun, MIAO Weipao. Fault diagnosis of rolling bearings based on variational mode decomposition and optimized support vector machine[J]. Journal of Chinese Society of Power Engineering. 2021, 41(3): 214-220+243 (in Chinese). [金江涛, 许子非, 李春, 缪维跑. 基于变分模态分解与优化支持向量机的滚动轴承故障诊断[J]. 动力工程学报, 2021, 41(3): 214-220+243.]
- [15] Zhu H, Cheng J, Zhang C, et al. Stacked pruning sparse denoising autoencoder based intelligent fault diagnosis of rolling bearings[J]. Applied Soft Computing, 2020, 88: 106060.
- [16] HU Chunsheng, LI Guoli, ZHAO Yong, CHENG Fangjuan. Summary of fault diagnosis methods for rolling bearings under variable working conditions[J/OL]. Computer Engineering and Applications: 1-23[2022-06-21].<http://kns.cnki.net/kcms/detail/11.2127.TP.20220523.1642.012.html> (in Chinese). [胡春生, 李国利, 赵勇, 成芳娟. 变工况滚动轴承故障诊断方法综述[J/OL]. 计算机工程与应用: 1-23[2022-06-21]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20220523.1642.012.html>.]
- [17] ZHAO Xiaoqiang, ZHANG Qingqing. Improved Alexnet based fault diagnosis method for rolling bearing under variable conditions, Journal of Vibration, Measurement & Diagnosis, 2020, 40(3): 472-480+623 (in Chinese). [赵小强, 张青青. 改进 Alexnet 的滚动轴承变工况故障诊断方法[J]. 振动. 测试与诊断, 2020, 40(3): 472-480+623.]
- [18] XIE Yuandong, LEI Wenping, HAN Jie, CHEN Lei. Research on bearing fault diagnosis using full vector RNN[J]. Machinery Design & Manufacture, 2021(9): 27-31 (in Chinese). [谢远东, 雷文平, 韩捷, 陈磊. 全矢 RNN 的轴承故障诊断研究[J]. 机械设计与制造, 2021(9): 27-31.]
- [19] Zou F, Zhang H, Sang S, et al. Bearing fault diagnosis based on combined multi-scale weighted entropy morphological filtering and bi-LSTM[J]. Applied Intelligence, 2021, 51(10): 6647-6664.
- [20] ZHANG Yonghong, ZHANG Zhongyang, WANG Lihua, SHAO Fan, LÜ Kaiyang. Bearing fault diagnosis method based on VAE-GAN and FLCNN unbalanced samples[J]. Journal of Vibration and Shock, 2022, 41(9): 199-209 (in Chinese). [张永宏, 张中洋, 赵晓平, 王丽华, 邵凡, 吕凯扬. 基于 VAE-GAN 和 FLCNN 的不均衡样本轴承故障诊断方法[J]. 振动与冲击, 2022, 41(9): 199-209.]
- [21] Hasan J, Islam M, Kim J. Acoustic Spectral Imaging and Transfer Learning for Reliable Bearing Fault Diagnosis under Variable Speed Conditions[J]. Measurement, 2019, 138: 620-631.
- [22] SHEN Tao, LI Shunming. CNN-LSTM method with batch normalization for rolling bearing fault diagnosis[J/OL]. Computer Integrated Manufacturing Systems: 1-16[2022-06-13]. <http://kns.cnki.net/kcms/detail/11.5946.TP.20210330.1708.003.html> (in Chinese). [沈涛, 李舜酩. 针对滚动轴承故障的批标准化 CNN-LSTM 诊断方法[J/OL]. 计算机集成制造系统: 1-16[2022-06-13]. <http://kns.cnki.net/kcms/detail/11.5946.TP.20210330.1708.003.html>.]
- [23] Stockwell G. Localization of the complex spectrum: the S transform[J]. IEEE Transactions on Signal Processing, 1996, 44(4): 998-1001.
- [24] Ruelle D. Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems[J]. Physica D Nonlinear Phenomena, 1992, 56(2_3): 185-187.
- [25] Fraser M, Swinney L. Independent coordinates for strange attractors from mutual information[J]. Physical Review A, 1986, 33(2): 1134-1140.
- [26] Kennel B, Brown R, Abarbanel I. Determining embedding dimension for phase-space reconstruction using a geometrical construction[J]. Physical Review A, 1992, 45(6): 3403-3411.
- [27] XU Yanwei, CAI Weiwei, XIE Tancheng, et al. Intelligent diagnosis of bearing faults of subway

traction motors based on information fusion under variable operating conditions[J]. Computer Integrated Manufacturing Systems, 2021,27(11): 12 (in Chinese). [徐彦伟, 蔡薇薇, 颀潭成,等. 变工况下基于信息融合的地铁牵引电机轴承故障智能诊断[J]. 计算机集成制造系统, 2021,27(11): 12.]

[28] LI Shan, TAN Jiwen, YU Kun. Composite fault diagnosis research of rolling bearing based on combination of neural network and improved D-S evidence theory[J]. Machine Tool & Hydraulics, 2018,46(1): 153-157+184 (in Chinese). [李 善, 谭继文, 俞 昆. 神经网络和改进 D-S 证据理论相结合的滚动轴承复合故障诊断研究[J]. 机床与液压, 2018,46(1): 153-157+184.]

[29] Althubiti S A, Alenezi F, Shitharth S, et al. Circuit Manufacturing Defect Detection Using VGG16 Convolutional Neural Networks[J]. Wireless Communications and Mobile Computing, 2022, 2022.

[30] Khorram A, Khalooei M, Rezghi M. End-to-end CNN+ LSTM deep learning approach for bearing fault diagnosis[J]. Applied Intelligence, 2021, 51(2): 736-751.

[31] Eren L, Ince T, Kiranyaz S. A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier[J]. Journal of Signal Processing Systems, 2019,91(2):179-189.

[32] Jiang H, Li X, Shao H, et al. Intelligent fault diagnosis of rolling bearings using an improved deep recurrent neural network[J]. Measurement Science and Technology, 2018,29(6):065107.

作者简介:

林诗麒(1997-), 女, 吉林白城人, 硕士研究生, 研究方向: 轴承故障诊断、机器学习、深度学习, E-mail:489502398@qq.com;

+陈智丽(1981-), 女, 辽宁鞍山人, 教授, 博士, 研究方向: 计算机视觉、模式识别、机器学习, E-mail: zzc@sjzu.edu.cn;

李宇鹏(1977-), 女, 云南昆明人, 二岗教授, 博士, 研究方向: 智能机器人、集成电路、计算机建模, E-mail: lyp@sjzu.edu.cn;

孟维迎(1987-), 男, 辽宁黑山县人, 副教授, 博士, 研究方向: 结构材料疲劳及可靠性研究、集成电路, E-mail:mengweiyong025@163.com。