

基于数字孪生与深度学习的轴承故障诊断方法研究

孙浩, 孙树峰*, 王津, 张兴波, 梁广磊, 王萍萍

(青岛理工大学 机械与汽车工程学院, 青岛 266525)

摘要:介绍了一种结合了数字孪生数据、阶比分析和深度残差收缩网络(DRSN)模型的滚动轴承故障诊断方法,其目的是解决现有轴承故障诊断研究中存在的变转速工况下诊断精度低的问题。该方法通过阶比分析提取变转速工况下振动信号的故障特征,随后将一维的故障特征通过格拉米角和场(GASF)转换为二维彩色图像,最后使用DRSN对轴承故障进行分类诊断。在数字孪生轴承故障诊断过程中,使用Unity 3D搭建可视化界面,实现轴承状态的实时远程监控。为验证所提出方法的有效性,使用渥太华大学轴承数据集和自测数据集进行故障诊断实验,结果表明,该方法具有较高的准确率。

关键词:故障诊断;数字孪生;深度学习;滚动轴承

中图分类号:TH133.33 **文献标志码:**A **文章编号:**1673-4602(2026)01-0060-08

DOI:10.3969/j.issn.1673-4602.2026.01.008

Research on the diagnosis method of bearing faults based on digital twin and deep learning

SUN Hao, SUN Shufeng*, WANG Jin, ZHANG Xingbo, LIANG Guanglei, WANG Pingping
(School of Mechanical and Automotive Engineering, Qingdao University of Technology, Qingdao 266525, China)

Abstract: A diagnosis method of rolling bearing faults combining digital twin data, order analysis and deep residual shrinkage network (DRSN) model is introduced in this study, which aims to address the challenge of low diagnosis accuracy under variable speed condition in the existing research on the diagnosis of bearing faults. The fault characteristics of vibration signals under variable speed conditions were extracted by order analysis. Subsequently, one-dimensional fault features were transformed into two-dimensional color images by Grameen Angle Summation Fields (GASF). Finally, DRSN was utilized for the classification and diagnosis of bearing faults. In the process of digital twin bearing fault diagnosis, a visualization interface was developed using Unity 3D, enabling real-time remote monitoring of bearing conditions. To verify the efficacy of the proposed method, fault diagnosis experiments were conducted using the bearing data set and self-test data set from the University of Ottawa, with results indicating a high level of accuracy of the method.

Key words: fault diagnosis; digital twin; deep learning; rolling bearings

收稿日期:2024-02-21

基金项目:山东省自然科学基金(ZR2023ME156);高等学校学科创新引智计划("111"计划)(D21017)

作者简介:孙浩(1999—),男,山东日照人。硕士,研究方向为数字孪生。E-mail:1260130886@qq.com。

*通信作者:孙树峰(1968—),男,山东诸城人。博士,教授,主要从事激光精密微纳制造技术的研究。

E-mail:sunshufeng@qut.edu.cn。

随着制造业的不断进步,从制造业大国向制造业强国的转型已成为我国经济发展的重大任务。而轴承作为旋转机械传动系统的核心,保证其良好的状态对于设备的稳定运行至关重要。然而,由于其自身具有故障隐蔽、不易拆卸等特点,轴承的故障维护一般是在故障发生的中后期进行,这会导致设备的可靠性降低、产生振动和噪声,甚至造成生命、财产损失。因此,轴承的故障检测对于设备的正常运行具有重要意义。

旋转机械故障诊断,即通过传感器数据对设备状态进行监测,一般来说诊断方法包括数据采集、特征提取、特征选择、故障分类 4 个步骤^[1]。在故障分类方面,深度学习因能够执行复杂识别任务的能力引起了广泛关注^[2]。相比于传统的故障诊断方法,深度学习驱动的诊断方法能够从故障信号中自动学习特征,从而具有更高的诊断精度^[3]。近年来,基于数据驱动的故障诊断技术不断涌现,如卷积神经网络(CNN)^[4]、深度置信网络^[5]、堆叠自动编码器^[6]、生成对抗网络^[7]等。其中,深度残差网络(ResNets)作为 CNN 的一种变体,因其能够解决梯度消失和梯度爆炸等问题而被广泛应用。ZHAO 等^[8]将 ResNets 和软阈值进行结合,提出了深度残差收缩网络(DRSN),将软阈值处理作为非线性转换层插入深度架构中,提高了对高噪声信号特征的学习能力,并实现了高精度的故障诊断。一些针对模型输入形式的研究证明,在进行分类任务时,以 CNN 架构为基础的深度学习网络在处理二维图像时具有更高的准确率^[9]。

作为一门拥有智能设计、智能建模、智能监测、智能预测等功能的新兴技术,数字孪生对工业数字化发展具有极大的推动作用^[10]。ZHANG 等^[11]提出了一种结合了数字孪生数据、迁移学习理论和增强的 ConvNext 深度学习网络模型的滚动轴承故障诊断框架。YANG 等^[12]结合虚实数据提出了一种数字孪生驱动的复合故障诊断方法,解决了水下生产控制系统复合故障的检测问题。XIA 等^[13]提出了一种基于数字孪生和深度迁移学习的机械智能故障诊断框架,通过孪生模型的生成数据训练模型。

为提升变转速工况条件下滚动轴承故障诊断的准确性,本文结合信号处理方法提取原始振动信号中的故障特征,使用深度学习模型对轴承故障进行诊断。因此,本文提出了基于数字孪生与深度学习的轴承故障诊断方法,通过阶比分析以及格拉米角和场(GASF)对轴承振动信号进行处理,然后使用 DRSN 模型进行故障分类并在 Unity 3D 中搭建可视化界面。

1 故障诊断方法总体框架

基于数字孪生的轴承故障诊断方法总体框架如图 1 所示,主要包含两大部分:虚拟环境以及实体环境。其中,虚拟环境包含基于 Unity 引擎的可视化显示以及数据驱动的轴承故障诊断模型两部分。基于传感器数据构建的孪生模型应具有在数字领域内精确复制物理实体的能力,从而创建包含多个维度和领域的全面孪生模型。因此,在实体环境中通过多种传感器实现全面地评估和监控。

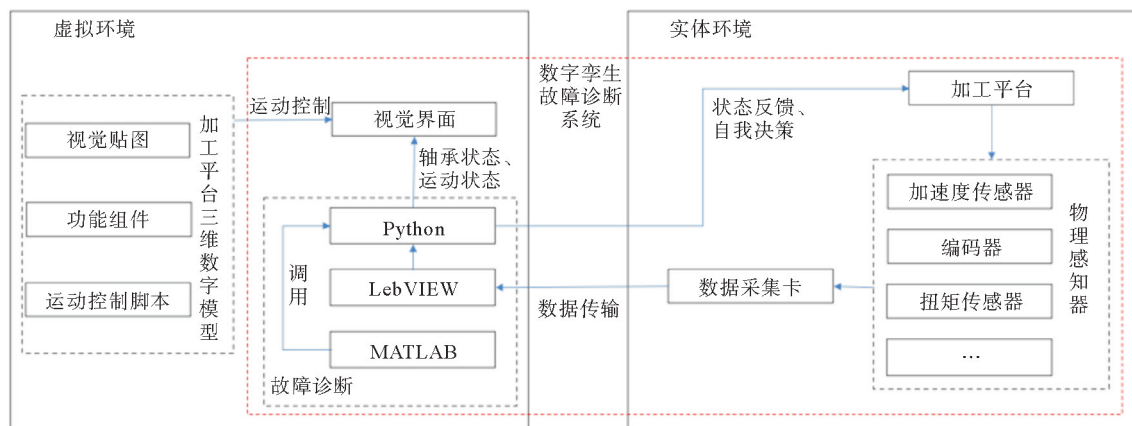


图 1 基于数字孪生的轴承故障诊断方法总体框架

2 轴承故障诊断方法原理

2.1 阶比分析算法

对于运行过程中存在较大转速波动的旋转机械,由于其转动频率发生变化,故障特征频率也会随之发

生变化。因此,对于变转速的振动信号,传统的分析方法不能有效识别故障。对于变转速状态下的旋转机械进行阶比分析的关键在于将非平稳的时域信号转换为相对平稳的角域信号。本文阶比分析实现步骤如下:

- 1) 采集振动信号和转速信号;
- 2) 将转速信号二值化处理,并根据结果计算脉冲个数和等角度重采样时刻;
- 3) 对振动信号进行快速傅里叶变换(FFT)并得到频谱图,并对其进行希尔伯特变换得到包络信号;
- 4) 将包络信号进行等角度重采样,然后再次进行 FFT,得到阶比谱。

2.2 深度学习模型

由 ZHAO 等^[8]在 ResNets 的基础上将软阈值作为非线性变换层插入到深度结构中,得到了能够有效消除噪声相关特征的 DRSN。与 ResNets 一样,DRSN 由卷积层(Conv)、一定数量的残差模块(RSBU)、批量归一化(BN)、激活函数层(ReLU)、全局平均池化层(GAP)和全连接层(FC)组成。

本文使用的 DRSN 整体架构如图 2 所示。其中,图形上方数字表示图像大小,下方数字表示通道数,模型中所使用的各种组件名称及参数在图片最下方列出。以 $(/2, 3, 1, K=C)$ 为例,其中 $/2$ 表示通过以 2 的步幅移动卷积核来减少特征映射的宽度, 3 表示卷积核宽度, 1 表示该组件的个数, K 为卷积核数, C 为通道数。将多个 RSBU 进行叠加,通过各种非线性变换来学习和判别特征,并以软阈值作为收缩函数,以消除噪声相关信息。

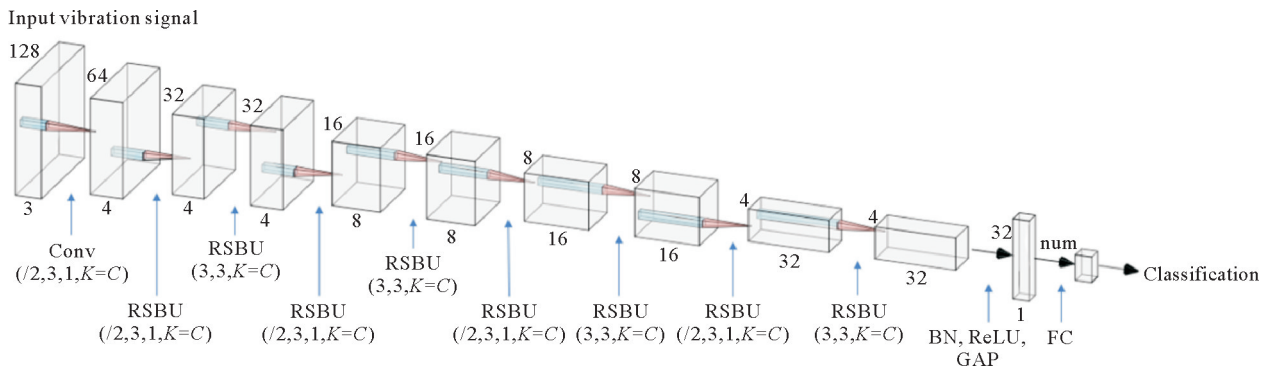


图 2 DRSN 整体架构

软阈值能够将噪声信号转换至接近零值的不重要区域,然后应用软阈值将特征中接近零的值转换为零。软阈值函数如式(1)所示:

$$y = \begin{cases} x - \tau, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ x + \tau, & x < -\tau \end{cases} \quad (1)$$

式中: x 为输入特征; y 为输出特征; τ 为阈值。

通过软阈值将负面特征调整为接近零的值,削弱无用特征影响的同时保留有用特征。导数定义如式(2)所示,输出相对于输入的导数是一个分段函数,取值为 1 或 0,有效解决了梯度消失和爆炸问题。

$$\frac{\partial y}{\partial x} = \begin{cases} 1, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ 1, & x < -\tau \end{cases} \quad (2)$$

在经典的信号去噪算法中,阈值的设置一直是一个难题,且最佳阈值会随实际情况变化。针对这一问题,DRSN 将单独的阈值应用于特征图的每个通道。具有通道阈值的 RSBU 如图 3 所示。

映射特征 x 经过绝对值运算和 GAP 层将其简化为一维矢量,然后传播到两层 FC 网络中。在 FC 网络中,神经元和特征图通道的数量相等。设置 FC 网络的输出范围在 $(0, 1)$,如式(3)所示:

$$\alpha_c = \frac{1}{1 + e^{-z_c}} \quad (3)$$

式中: α_c 为第 c 个缩放参数; z_c 为第 c 个神经元处的特征。

阈值由式(4)确定:

$$\tau_w = \alpha_w \cdot \text{average} |x_{p,q,w}| \quad (4)$$

式中: τ_w 为特征图的第 w 个通道的阈值; p 、 q 和 w 分别为特征图的宽度、高度和通道。

2.3 格拉米角和场

本文使用格拉米角和场将阶比分析结果编码为二维彩色图像。该方法能够保持一维信号与格拉米矩阵中阶比序列的相关性,其中每个元素为对应角度三角函数的值^[14]。首先对原始时间序列进行归一化操作,定义在式(5)中:

$$\tilde{\epsilon}_i = \frac{\epsilon_i - \epsilon_{\min}}{\epsilon_{\max} - \epsilon_{\min}} \quad (5)$$

式中: $\tilde{\epsilon}_i$ 为归一化操作后得到的信号; ϵ_i 为对应时刻 i 的振动信号; ϵ_{\max} 为 ϵ 的最大值; ϵ_{\min} 为 ϵ 的最小值。

使用极坐标来表示归一化的时间序列 $\tilde{\epsilon}$,时间戳作为半径 r ,如式(6)所示:

$$\begin{cases} \phi_i = \arccos \tilde{\epsilon}_i, & 0 \leq \tilde{\epsilon}_i \leq 1, \tilde{\epsilon}_i \in \tilde{\Sigma} \\ r = t_i, & t_i \in N \end{cases} \quad (6)$$

式中: ϕ_i 为第 $i-1$ 个序列的角度值; $\tilde{\Sigma}$ 为振动信号的时间序列; t_i 为当前的时间戳代码。

将新的时间序列转换为极坐标系后,通过每个点之间的三角和,使用角度视角来识别不同时间间隔中的时间相关性,由式(7)定义:

$$\mathbf{G} = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \cdots & \cos(\phi_1 + \phi_n) \\ \cos(\phi_2 + \phi_1) & \cdots & \cos(\phi_2 + \phi_n) \\ \vdots & \cos(\phi_i + \phi_i) & \vdots \\ \cos(\phi_n + \phi_1) & \cdots & \cos(\phi_n + \phi_n) \end{bmatrix} \quad (7)$$

可以看出,在这种编码方式下,二维图像保留了从左上到右下的时间顺序。振动信号的信息被记录在主对角线位置,其他位置记录不同信号之间的关系。当原始信号长度为 n 时,GASF 编码得到的数值矩阵 \mathbf{G} 大小为 $n \times n$ 。

3 故障诊断模型实验验证

3.1 实验设备与模型训练配置

渥太华大学在机械故障模拟器中采集了变转速工况下的轴承振动数据,实验装置如图 4^[15]所示。轴上安装有 2 个轴承,其详细参数见表 1,左侧为健康轴承,右侧为实验轴承。加速度计安装在实验轴承座上,转速由增量编码器测量。

数据集包含 36 个 .mat 文件,包括轴承的加速、减速、先加速再减速和先减速再加速 4 种工况,以及健康、内圈故障和外圈故障 3 种轴承状态,共 12 种不同情况,每种情况采集 3 个文件。每个文件中包含 2 个通道,通道 1 保存振动信号,通道 2 保存转速信号。信号的采样频率为 20 000 Hz,每个文件的采样时长为 10 s。

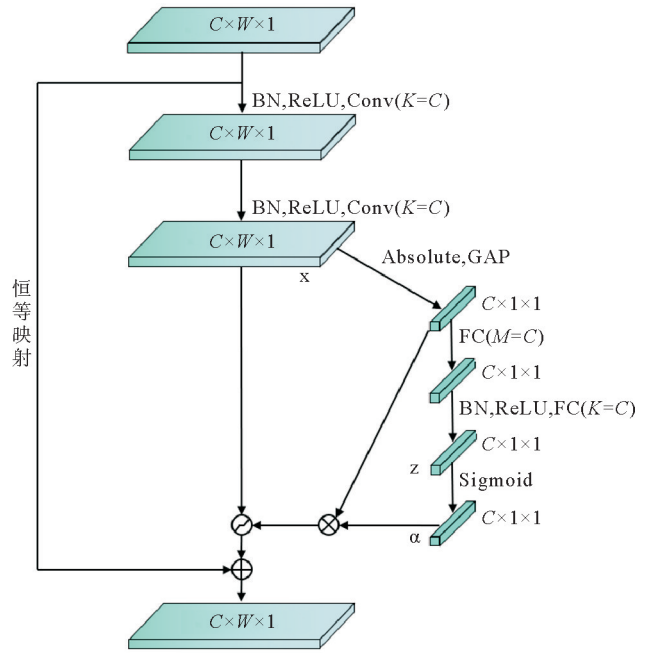


图 3 RSBU 结构

本文所有故障诊断实验均在 Python 3.6 编码环境下进行,深度学习框架在 TensorFlow1.x 中构建,阶比分析在 MATLAB 2023a 中进行。测试在 1 台配备 AMD Ryzen 9 5900HX CPU 3.30 GHz、16 GB 内存、NVIDIA GeForce RTX 3060 Laptop GPU 和 Windows 11、64 位操作系统的计算机上进行。

3.2 数据集制作

将 3.1 中的 36 个原始数据文件导入 MATLAB 中进行阶比分析,每个 .mat 文件中包含 2 个通道,“Channel_1”为振动加速度值,“Channel_2”为转速脉冲。首先使用低通滤波器过滤 5000 Hz 以上的高频噪声,并按照每转脉冲数(1024)将上述 2 个变量中的信号进行分割,最后按照 2.1 中的阶比分析步骤对分割后的数据进行阶比分析。结果如图 5 所示,可以在故障阶次及其倍频上看到明显的幅值。

本文使用 128×128 像素图像作为 DRSN 模型的输入,为使阶比分析与图像大小匹配,使用线性插值函数对一维阶比序列的长度进行缩减以适应图像大小。通过 GASF 对阶比分析结果进行编码得到的图像如图 6 所示。图像数据集中包含 3 种轴承状态各 250 张图片,共 750 张图片,格式为 .JPG。

3.3 故障分类实验

由于本文所使用的图像为 128×128 像素,因此模型的输入大小设置为 $3 \times 128 \times 128$ 。数据集按照 8:2 的比例划分为训练集和预测集,训练过程中将训练集再次按照 8:2 的比例划分,用于训练和测试。模型的训练采用 5 折交叉验证来防止过拟合,将训练集分为 5 份,每折训练使用其中 1 个子集作为验证集,其余 4 个子集用于训练。样本使用 train_test_split 函数打乱并划分,random_state 设置为 42。本文中模型的训练采用的训练轮数为 40 轮,步长为 32。

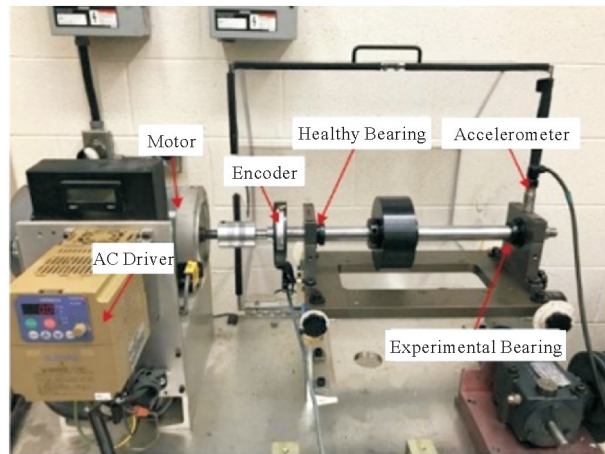


图 4 渥太华大学实验装置^[15]

表 1 渥太华大学轴承参数

轴承型号	节圆直径/mm	滚动体直径/mm	滚动体个数/个
ER16K	38.52	7.94	9

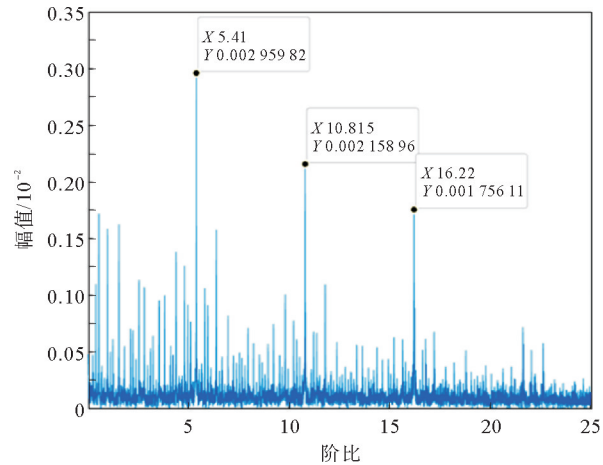


图 5 阶比分析结果

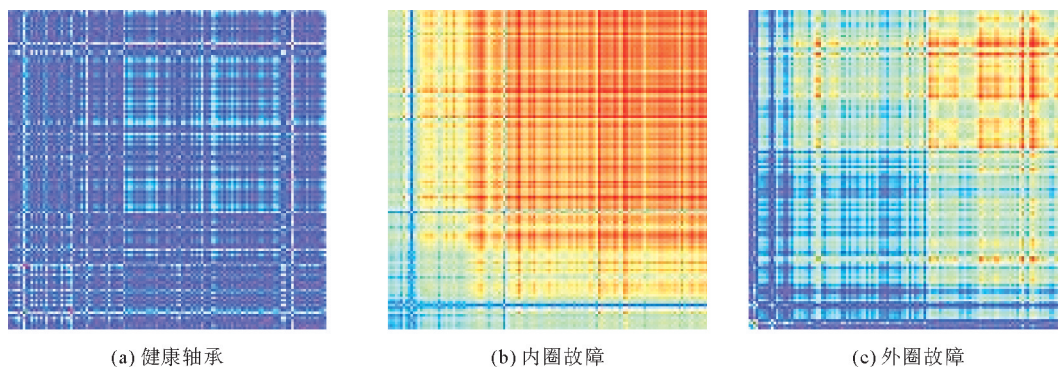


图 6 阶比分析结果图像

分别将阶比分析后的二维图像(i)、原始信号的二维图像(ii)以及一维原始信号(iii)作为 DRSN 的输入,并对训练过程中的准确率和损失值变化进行记录,如图 7 所示。可见,输入(i)的模型在较短时间内趋于稳定,准确率在前 100 步内快速提高,在 200 步左右逐渐收敛;损失值在前 500 步内快速下降,在 1000 步之后逐渐收敛。而输入(ii)和(iii)的准确率分别在 400 步和 1000 步收敛,损失值在 1000 步和 1500 步收敛。因此,当使用(i)作为模型输入时,训练过程中的准确率和损失值收敛速度更快,且损失值降低到了更低的水平。

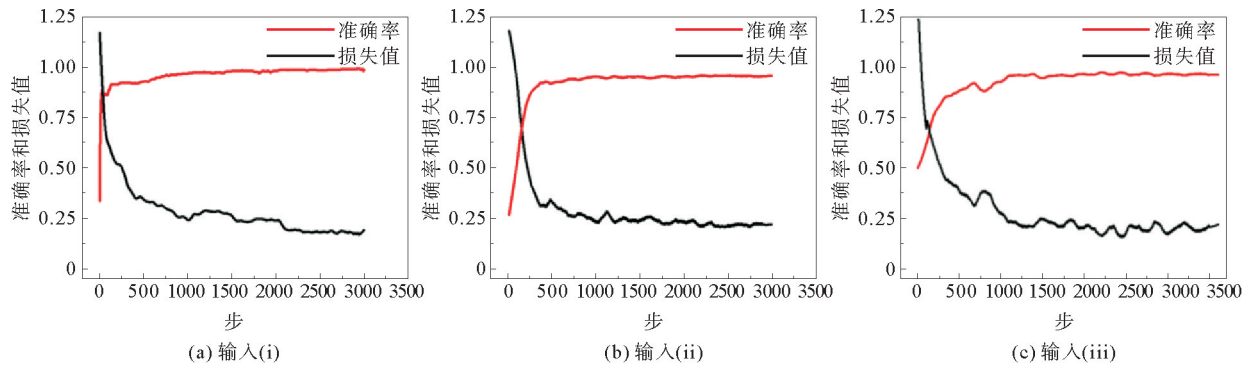


图 7 模型训练准确率和损失值变化

3.4 轴承故障预测结果

上述 3 个数据集输入模型得到的平均准确率和平均耗时如表 2 所示,表中数据均为预测 5 次的平均值。由表 2 可知,输入(i)的预测准确率最高,且平均耗时仅比最快的输入(iii)慢 0.267 s,能够满足对于响应时间的要求。输入(i)的轴承故障预测结果如图 8 所示,其中横轴为预测标签,纵轴为真实标签,数字为样本数量。其中,H 表示健康轴承,I 表示内圈故障,O 表示外圈故障。图 8 中预测结果为训练 5 次取最佳,最佳预测准确率为 98.67%。

表 2 模型训练结果

模型输入	平均准确率/%	平均耗时/s
(i)	98.13	1.270
(ii)	91.11	1.070
(iii)	83.70	1.003

4 诊断结果可视化界面的搭建与测试

本文采用 Unity3D 引擎对轴承故障诊断系统的可视化界面进行开发,其中可视化映射采用 UGUI 组件实现。所开发的可视化界面如图 9 所示,其中主要包括轴承数据显示模块、轴承故障诊断模块、设备显示模块。轴承数据显示模块将振动数据和转速数据以折线图的形式进行可视化,将轴承的工作状态进行直观展示;轴承故障诊断模块对轴承状态进行分类预测,并在轴承放大图中进行展示;设备显示模块展示设备的实时运动状态。

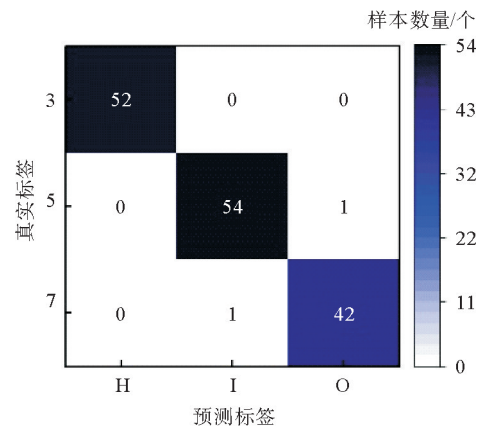


图 8 输入(i)的故障预测结果

轴承故障诊断模块作为可视化界面的核心,由数据处理和深度学习模型两部分组成。数据处理部分在 MATLAB 中进行,其中包括数据分割、滤波以及阶比分析。深度学习模型部分使用 Python 进行,其中包括 MATLAB 函数调用、GASF 二维图像编码、深度学习模型训练及模型权重保存、与 C# 建立 socket 通信输出预测结果。

本文根据传感器数据对 Unity 中的三维模型进行运动控制,传感器数据通过 National Instruments 采集卡传输至 LabVIEW 2018。测试开始后,使用 UDP 通信建立 Python 与 LabVIEW 间的数据连接,在 Python 中调用 MATLAB 程序文件对分割后的原始振动数据进行阶比分析,将阶比分析结果进行二维图

像编码后输入深度学习模型进行训练,最后将训练过程中的准确率、损失值以及最终的预测结果导出并在 Unity 中进行可视化。模型训练过程和结果如图 10 所示,图中 IO 为内外圈混合故障。训练 5 次的平均准确率为 93.47%,图 10(b)的准确率最高值为 93.75%,平均消耗时间为 1.082 s。

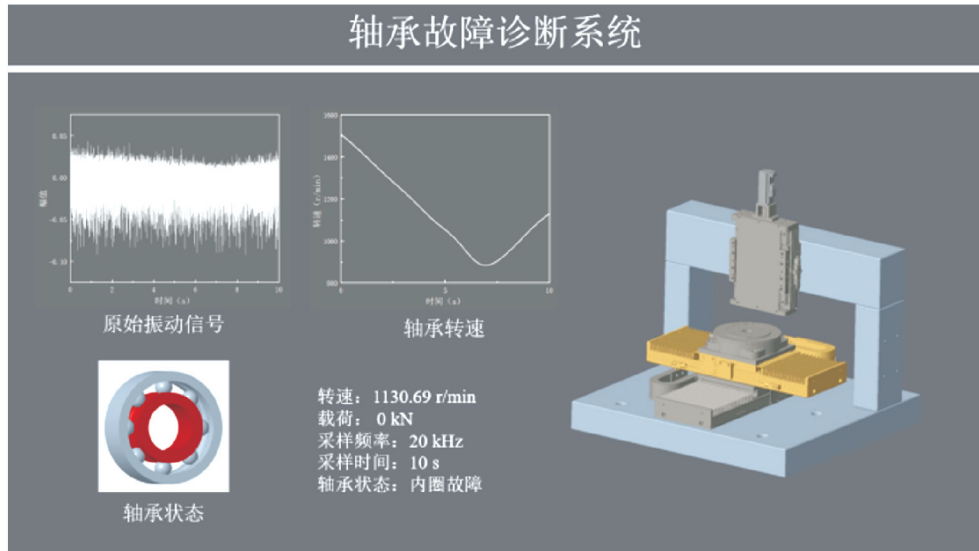


图 9 轴承故障诊断结果可视化界面

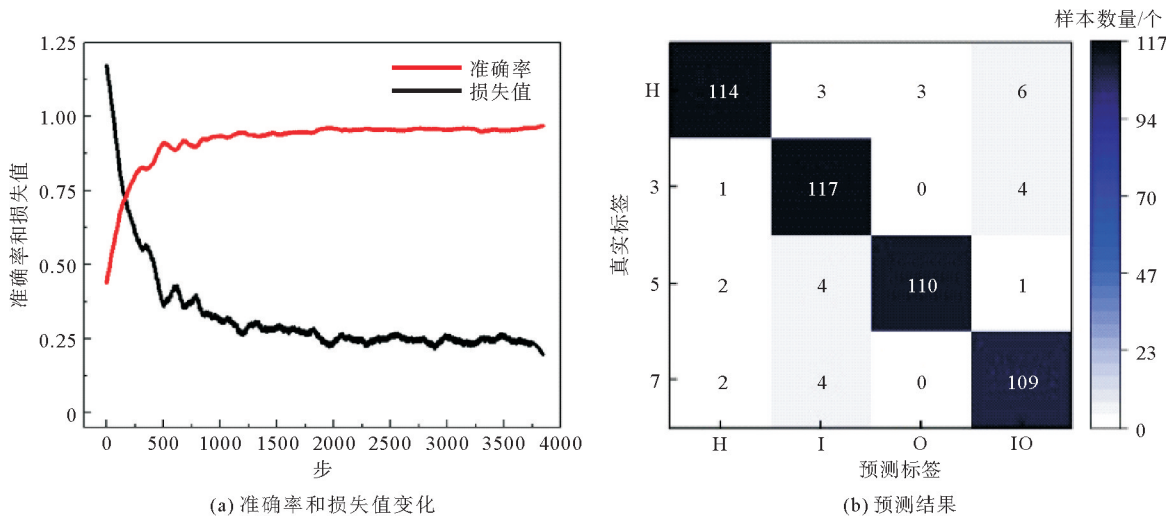


图 10 自测数据集故障诊断结果

5 结论

本文基于 DRSN 模型和阶比分析算法提出了基于数字孪生的轴承故障诊断方法。使用传感器采集的振动和转速数据,并在 Python 中对其进行二维图像编码及深度学习故障诊断。对于变转速工况下的振动信号,使用阶比分析算法将非稳态的振动信号转换为稳态信号,从而提取出明显的故障特征。可视化界面在 Unity3D 中构建,使用 C# 脚本控制三维模型运动,Python 中的故障诊断结果通过 Socket 协议传输到 C# 并实现可视化。结论如下:

- 1) 使用渥太华大学轴承数据集进行对比实验,在相同的模型参数和结构条件下,相比于使用原始信号,通过阶比分析提取的故障特征能够有效提升故障诊断的准确率。
- 2) 通过引入格拉米角和场,将一维信号编码为二维彩色图像,能够进一步发挥 CNN 架构模型的图像处理性能。

3) 通过实现传感器—数据处理—故障诊断—结果可视化的数据连接,基于数字孪生和深度学习的轴承故障诊断方法能够对轴承的实时状态进行监测。

该方法在诊断速度及准确率方面较为理想,符合基本的设计要求,能够满足实际应用中轴承故障诊断的需求。但是,所构建的系统还存在一定的缺陷,首先,对于数据集中存在的异常值,本文方法并不能有效处理,仍需要进一步的数据处理以提高准确率;其次,需要较长时间的稳定运行数据以支持故障诊断,短时的反复变转速运动数据诊断结果较差。

参考文献(References):

- [1] MUSHTAQ S, ISLAM M, SOHAIB M. Deep learning aided data-driven fault diagnosis of rotatory machine: A comprehensive review[J]. *Energies*, 2021, 14(16): 5150-5150.
- [2] YANG Y Y, HAQUE M M M, BAI D L, et al. Fault diagnosis of electric motors using deep learning algorithms and its application: A review[J]. *Energies*, 2021, 14(21): 7017-7043.
- [3] 王婧, 许志伟, 刘文静, 等. 滚动轴承健康智能监测和故障诊断机制研究综述[J]. *计算机科学与探索*, 2024, 18(4): 878-898.
WANG Jing, XU Zhiwei, LIU Wenjing, et al. Review of research on rolling bearing health intelligent monitoring and fault diagnosis mechanism[J]. *Journal of Frontiers of Computer Science and Technology*, 2024, 18(4): 878-898.
- [4] TAO H, JIA P, WANG X, et al. Real-time fault diagnosis for hydraulic system based on multi-sensor convolutional neural network[J]. *Sensors*, 2024, 24(2): 353-374.
- [5] MIAO X, QUAN H, CHENG X, et al. Fault diagnosis of oil-immersed transformers based on the improved neighborhood rough set and deep belief network[J]. *Electronics*, 2023, 13(1): 5-19.
- [6] 李兵, 梁舒奇, 单万宁, 等. 基于改进正余弦算法优化堆叠降噪自动编码器的电机轴承故障诊断[J]. *电工技术学报*, 2022, 37(16): 4084-4093.
LI Bing, LIANG Shuqi, SHAN Wanning, et al. Motor bearing fault diagnosis based on improved sine and cosine algorithm for stacked denoising autoencoders[J]. *Transactions of China Electrotechnical Society*, 2022, 37(16): 4084-4093.
- [7] ZHANG C, PENG K, DONG J, et al. A novel exergy-related fault detection and diagnosis framework with transformer-based conditional generative adversarial networks for hot strip mill process[J]. *Control Engineering Practice*, 2024, 144: 105820-105839.
- [8] ZHAO M, ZHONG S, FU X, et al. Deep residual shrinkage networks for fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2019, 16(7): 4681-4690.
- [9] CUI J, ZHONG Q, ZHENG S, et al. A lightweight model for bearing fault diagnosis based on gramian angular field and coordinate attention[J]. *Machines*, 2022, 10(4): 282-305.
- [10] PENG F, ZHENG L, PENG Y, et al. Digital twin for rolling bearings: A review of current simulation and PHM techniques[J]. *Measurement*, 2022, 201(7): 111728-111748.
- [11] ZHANG C, QIN F, ZHAO W, et al. Research on rolling bearing fault diagnosis based on digital twin data and improved ConvNext[J]. *Sensors*, 2023, 23(11): 5334-5356.
- [12] YANG C, CAI B, WU Q, et al. Digital twin-driven fault diagnosis method for composite faults by combining virtual and real data[J]. *Journal of Industrial Information Integration*, 2023, 33(3): 100469-100479.
- [13] XIA M, SHAO H, WILLIAMS D, et al. Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning[J]. *Reliability Engineering & System Safety*, 2021, 215(1/2): 107938-107947.
- [14] HUANG H, BADDOUR N. Bearing vibration data collected under time-varying rotational speed conditions[J]. *Data in Brief*, 2018, 21(1): 1745-1749.
- [15] WANG Z, OATES T. Imaging time-series to improve classification and imputation[C]//IJCAI'15: Proceedings of the 24th International Conference on Artificial Intelligence, Buenos Aires, 2015: 3939-3945.

(责任编辑 张晓靖;英文校审 程文华)