

基于 Ce-ATRP-CNN 的滚动轴承故障诊断方法

蒋丽英, 盛歆然, 李贺

(沈阳航空航天大学 自动化学院, 沈阳 110136)

摘要: 针对滚动轴承故障振动信号在实际运行环境下故障特征易被噪声影响, 一维信号转换为二维图像时不能有效将信号的故障特征反映到图像中, 从而影响故障识别的问题, 创新性地将自适应阈值融入递归图编码过程, 提出了一种名为自适应阈值递归图编码 (ATRP) 的图像转换方法; 在此基础上, 进一步构建了 Ce-ATRP-CNN 故障诊断模型; 通过倒谱分析对原始振动信号预处理, 可以一定程度地滤除噪声; 应用提出的 ATRP 将处理后的一维振动信号转换为二维故障特征图像—ATRP 图; 将 ATRP 图作为卷积神经网络的输入实现对滚动轴承故障特征的深入提取和识别诊断; 应用 CWRU 和 HUST 轴承故障数据集进行了试验验证, 故障诊断准确率分别达到 99.17% 和 99.05%; 实验结果表明该模型具有良好的故障识别能力和泛化性。

关键词: 故障诊断; 滚动轴承; 故障特征提取; 图像编码; 卷积神经网络

A Rolling Bearing Fault Diagnosis Method Based on Ce-ATRP-CNN

JIANG Liying, SHENG Xinran, LI He

(College of Automation, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: The fault features of vibration signals from rolling bearing are easily affected by noises in actual operating environments, which are ineffectively reflect in the images with an one-dimensional signal converted into an two-dimensional image, thus affecting the identification of faults. To address these issues, by innovatively introducing the adaptive threshold into the recurrence plot encoding, an image conversion method for the adaptive threshold recurrence plot (ATRP) encoding is proposed. On this basis, a Ce-ATRP-CNN fault diagnosis model is further presented. The original vibration signal is pre-processed through cepstrum analysis, which can filter out the noise to a certain extent. The proposed ATRP converts the processed one-dimensional vibration signal into a two-dimensional fault feature image, namely ATRP image. As the input of the convolutional neural network (CNN), the ATRP image is used to achieve in-depth extraction and diagnostic identification for the fault features of rolling bearings. Experiments are verified on the CWRU and HUST bearing fault datasets, and the accuracies of the fault diagnosis reach up to 99.17% and 99.05%, respectively, which shows that the model has a good fault recognition ability and generalization performance.

Keywords: fault diagnosis; rolling bearing; fault feature extraction; image encoding; CNN

0 引言

滚动轴承作为旋转机械的重要部件, 其状态直接影响整个机械设备的安全运行^[1]。滚动轴承一旦发生故障, 通过准确有效地诊断及时发现故障并采取相应的维护措施, 对于机械设备的持续稳定运行至关重要。

滚动轴承故障诊断方法可分为以下 3 种技术路线: 基于信号处理的故障诊断、基于传统机器学习的故障诊断和基于深度学习的故障诊断。基于信号处理的方法主

要包括时域分析、频域分析和时频分析。文献 [2] 采用灰狼算法对变分模态分解 (VMD, variational mode decomposition) 参数进行优化, 并对信号进行分解, 计算故障信号的时域参数、能量、排列熵提取故障特征。对于非线性非平稳信号, 特征选择 (如均值、频谱峰值、小波系数等) 容易导致关键信息丢失影响诊断精度。基于机器学习的方法主要在信号处理的基础上引入支持向量机 (SVM, support vector machine)、决策树等算法将故障特征分类。文献 [3] 把 VMD 能量熵和

收稿日期: 2025-03-28; 修回日期: 2025-05-21。

基金项目: 国家自然科学基金项目 (62003223)。

作者简介: 蒋丽英 (1976-), 女, 博士, 副教授, 硕士生导师。

引用格式: 蒋丽英, 盛歆然, 李贺. 基于 Ce-ATRP-CNN 的滚动轴承故障诊断方法 [J]. 计算机测量与控制, 2026, 34(4): 19-26.

各 IMF 分量的能量比作为特征输入 SVM 中进行故障识别。但机器学习在复杂工况和故障类型较多的情况下步骤较复杂且诊断效果欠佳。

目前,深度学习已经广泛应用于滚动轴承故障诊断研究中^[4]。由于振动信号往往是一维时间序列,因此如何将一维信号转换为二维图像是应用深度学习方法的关键问题^[5]。早期的图像转换方法主要包括灰度图、小波变换、短时傅里叶变换(STFT, short time fourier transform)^[6]等。文献[7]提出将信号直接转换为灰度图的方法,将灰度图输入卷积神经网络(CNN, convolutional neural network)模型中进行故障识别,但将信号直接转换为灰度图的过程中噪声可能会被放大或引入新的噪声成分,从而影响识别准确性。文献[8]提出将一维轴承数据通过小波变换转化为二维时频图输入残差网络中进行故障诊断。小波变换过程中小波基函数的选择以及分解层数的选择都会影响关键故障特征的提取。文献[9]将振动信号通过 STFT 得到时频谱样本输入卷积神经网络进行故障诊断。STFT 假设信号在每个时间窗内是平稳的,然而实际的振动信号往往是非平稳的,这会导致 STFT 无法准确反映信号的时变特性。

近几年,随着故障特征提取技术和图像处理技术的发展,马尔可夫转移场(MTF, markov transition fields)^[10]、递归图(RP, recurrence plot)^[11]、对称点图谱^[12]等图像转换方法被应用到故障诊断领域中。文献[13]提出利用 MTF 将一维信号转换为二维特征图后输入到含有注意力机制的多尺度卷积神经网络进行诊断。虽然 MTF 可以将信号转换为具有时间相关性的二维图像,但噪声在信号到图像的转换过程中会被 MTF 放大(如相邻时间窗的随机波动被误判为状态转移),降低特征图像的可区分性。文献[14]提出了递归图编码技术与残差网络的滚动轴承故障诊断模型。虽然可以通过重复出现的二维表示可视化和研究多维相空间轨迹,但是其嵌入维数、延迟时间和距离阈值的选择会影响图像中故障特征呈现。对称点图谱^[15]是基于极坐标转换的方式将振动信号转换为二维雪花图,其转换过程中对称性偏转角、放大系数和时间间隔系数的参数选择都会影响特征图的构建^[16]。

针对上述问题,本文提出了一种自适应阈值递归图(ATRP, adaptive threshold recurrence plot)的故障特征图像转换方法,旨在通过构建自适应阈值对递归图进行编码,以削弱噪声的影响来突出信号的冲击性成分。同时,构建基于 Ce-ATRP-CNN 的滚动轴承故障诊断模型:首先,采用倒谱分析(Ce, cepstrum)对原始故障振动信号预处理。然后,应用 ATRP 图像转换方法,针对预处理后的一维振动信号采用自适应阈值对递归图

编码矩阵筛选后生成二维故障特征图像,在一定程度上消除噪声对故障特征图像的影响从而更加突出信号的冲击性成分,以此突出故障特征。最后,将 ATRP 图像作为 CNN 网络的输入,从二维特征的角度实现滚动轴承故障特征的深入提取和故障的识别诊断。实验结果表明本文模型可以较为准确地地区分不同滚动轴承故障类别及故障程度,采用结构较为简单的 CNN 网络模型就可以得到较高的故障诊断准确率;在添加弱噪声的条件下本文方法仍能取得良好的故障诊断准确率。

1 自适应阈值递归图

针对递归图编码存在噪声点间距离值编码会掩盖图像中故障冲击性特征的问题,本文提出自适应阈值递归图编码方法。本方法在递归图矩阵 D 基础上,引入自适应阈值 δ 对递归图进行编码,提取 D 阵中表征故障冲击特性的距离元素,从而突出二维图像中的冲击性纹理特征。

1.1 递归图

递归图是一种用于可视化和分析时间序列或动态系统的方法。它将时间序列转化为图形化的表示形式,以便分析时间序列中的重复模式和结构。

设 $T = [T_1, T_2, \dots, T_n]$ 为包含 n 个点的时间序列,则序列 T 的递归图 $D \in R^{n \times n}$ 的计算表达式为:

$$D(i, j) = d_{ij} = f(T_i, T_j) \quad (1)$$

$$i, j = 1, 2, \dots, n$$

式中, $f(\cdot)$ 为距离函数, 3 种不同实现形式如下^[17]:

1) 绝对距离: 用于衡量两个时间序列点 T_i 和 T_j 之间的绝对差值:

$$f(T_i, T_j) = |T_i - T_j| \quad (2)$$

2) 相对距离: 通过计算两个时间序列点的比值来衡量它们之间的相对差异:

$$f(T_i, T_j) = \frac{T_i}{T_j} \quad (3)$$

3) 乘积距离: 可以反映两个时间序列点之间的联合效应或相互作用:

$$f(T_i, T_j) = T_i \times T_j \quad (4)$$

公式(2)~(4)对应3种不同的距离度量方式:绝对距离对信号幅值变化较为敏感,当故障导致振动信号的幅值发生明显变化时,能够有效地捕捉到这种差异;相对距离在处理近零值的数据时可能会出现问题,它对信号的绝对幅值变化不敏感;乘积距离是一种在多元变量数据分析中用于衡量两个向量之间差异的方法,尤其适用于当故障或异常导致信号多个成分同时发生变化时捕捉这种联合变化的情况。综上,本文使用绝对距离计算时间序列中每两个点之间的距离,编码时间序列 T 状态的递归程度。

1.2 自适应阈值

本文依据 1.1 节递归图矩阵 \mathbf{D} 和初始化百分位 $\alpha\%$ 确定自适应阈值 δ , 其具体计算步骤如下:

1) 对矩阵 \mathbf{D} 中的元素按从小到大的顺序进行排序, 构成一维向量 $\mathbf{D}_s \in R^{1 \times n^2}$, 即:

$$\mathbf{D}_s = \text{sort}(d_{ij}) \quad (5)$$

式中, $\text{sort}(\cdot)$ 表示将矩阵元素按从小到大的顺序排列成一维的向量。

2) 按照给定的百分位 $\alpha\%$, 确定位置索引 γ : 向量 \mathbf{D}_s 的前 $\alpha\%$ 元素中最大元素对应的位置作为索引 γ , 即:

$$\gamma = \lceil \alpha \times n^2 / 100 \rceil \quad (6)$$

式中, $\lceil \cdot \rceil$ 表示向上取整。

3) 确定自适应阈值 δ . 取向量 \mathbf{D}_s 中第 γ 个元素值作为阈值 δ , 即有:

$$\delta = \mathbf{D}_s(\gamma) \quad (7)$$

由式 (6) 可知, α 的大小决定了位置索引 γ 的大小。对于相同的矩阵 \mathbf{D} , α 值越大, 则 γ 就越大, 进而自适应阈值 δ 也就越大; 反之, α 值越小, 自适应阈值 δ 越小。通常取 $\alpha > 85$ 时编码的故障特征图即可可视化出不同故障类别的差异。

1.3 自适应阈值递归图

将 1.2 节所提出的自适应阈值与递归图相结合, 通过比较递归图矩阵 \mathbf{D} 中所有元素与自适应阈值的大小, 对矩阵 \mathbf{D} 中元素进一步筛选, 构建编码矩阵 \mathbf{R} , 简称 ATRP 图。

对于任意长度为 n 的时间序列 T , 自适应阈值递归图构建过程如下:

1) 构建递归图矩阵 \mathbf{D} . 根据 1.1 节中的绝对距离公式计算振动信号时序状态点间的距离值 d_{ij} , 从而构建递归图矩阵 \mathbf{D} :

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix} \quad (8)$$

$$d_{ij} = |T_i - T_j| \quad i, j = 1, 2, \dots, n \quad (9)$$

式中, T_i 和 T_j 分别表示振动信号中第 i 和 j 个时间序列点的幅值。

2) 确定自适应阈值 δ . 根据公式 (5) ~ (7) 计算自适应阈值 δ ;

3) 构建 ATRP 图像编码矩阵 \mathbf{R} . 将矩阵 \mathbf{D} 中所有元素与自适应阈值 δ 相比较, 若 $d_{ij} \leq \delta$, 则在编码矩阵 \mathbf{R} 中的对应元素 r_{ij} 置 1; 反之 r_{ij} 置 0, 即:

$$r_{ij} = \begin{cases} 1 & d_{ij} \leq \delta \\ 0 & d_{ij} > \delta \end{cases} \quad (10)$$

因此, 构建的编码矩阵 \mathbf{R} 表示为:

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix} \quad (11)$$

2 基于 Ce-ATRP-CNN 的滚动轴承故障诊断方法

基于 Ce-ATRP-CNN 的滚动轴承故障诊断结构如图 1 所示。首先, 使用倒谱分析^[18]对原始振动信号进行预处理; 其次, 采用本文提出的 ATRP 方法将处理后的一维信号转换为故障特征图像, 即 ATRP 图; 最后, 将 ATRP 图作为 2D-CNN 故障诊断网络的输入, 实现对滚动轴承故障特征的深入提取和故障识别。

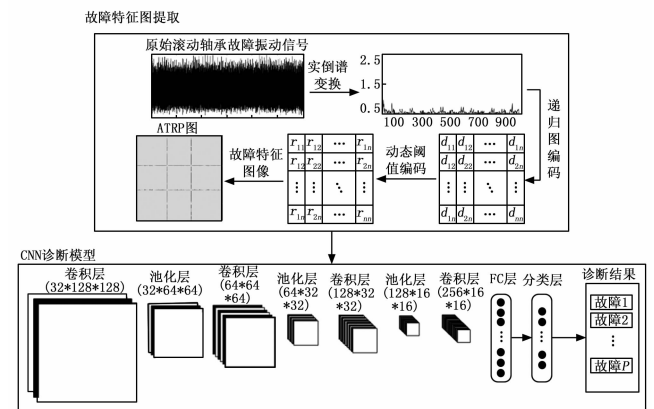


图 1 Ce-ATRP-CNN 诊断模型结构

2.1 Ce-ATRP-CNN 模型训练流程

训练 Ce-ATRP-CNN 故障诊断模型具体流程如下:

1) 构建训练样本集。假设共 P 种故障类型。若取样本长度为 L , 样本数为 N , 则第 P 种故障的样本集记为 $X_p = \{x_i^p\}$. 其中, $x_i^p \in R^L$ 为训练样本集中第 p 种故障的第 i 个样本 ($i = 1, 2, \dots, N; p = 1, 2, \dots, P$);

2) 倒谱变换。倒谱分析关注于频谱中的周期性冲击成分, 可以降低噪声对振动信号的影响。因此, 本文采用倒谱变换对原始振动信号进行预处理。样本集中各样本 x_i^p 的倒谱变换可表示为:

$$\tilde{x}_i^p = F^{-1}[\log_{10} F(x_i^p)] \quad (12)$$

式中, $F(\cdot)$ 和 $F^{-1}[\cdot]$ 分别为傅里叶变换和傅里叶逆变换; $\log_{10}(\cdot)$ 为对数谱。

3) 构建 ATRP 图集 R_p . 应用 1.3 节 ATRP 图像转换方法, 将样本 \tilde{x}_i^p 转换为编码矩阵 $\mathbf{R}_i^p \in \mathbb{R}^{L \times L}$, 并按尺度 L' 对 \mathbf{R}_i^p 进行缩放处理得到 $\tilde{\mathbf{R}}_i^p \in \mathbb{R}^{L' \times L'}$, 以保证输入图像的宽度和高度符合模型输入要求, 从而得到训练集 $\mathbf{R}_p = \{\tilde{\mathbf{R}}_i^p\}$.

4) 构建 CNN 故障诊断网络模型。根据数据类型、

输入图像大小以及故障类别, 设置 CNN 模型的初始网络参数: 定义神经网络架构, 包括卷积层和卷积核的数量、池化层和全连接层的选择等^[19]。其中, R_p 为网络模型的输入, 输出共 P 个节点, 分别对应相应的故障类型。

5) 训练模型。设置合适的迭代次数, 遍历整个训练集。在每一轮训练中, 使用损失函数计算预测结果与实际标签之间的损失, 根据损失计算梯度, 并通过反向传播更新网络模型参数, 直到损失不再下降, 从而得到一个训练好的 CNN 网络模型, 即为故障诊断模型, 使用该模型对滚动轴承进行故障诊断。

2.2 Ce-ATRP-CNN 模型测试诊断流程

应用 Ce-ATRP-CNN 模型对滚动轴承进行故障诊断流程如下:

- 1) 采集滚动轴承故障振动信号, 构建待诊断的样本 $x_{new} \in \mathbb{R}^{1 \times L}$;
- 2) 根据公式 (12) 对 x_{new} 进行倒谱变换预处理, 得到处理后的样本为 \tilde{x}_{new} ;
- 3) 根据 ATRP 图像转换方法, 将 \tilde{x}_{new} 转换为编码矩阵 $R_{new} \in \mathbb{R}^{L \times L}$, 并按尺度 L' 对 R_{new} 进行缩放处理得到 $\tilde{R}_{new} \in \mathbb{R}^{L' \times L'}$;
- 4) 应用训练好的 CNN 模型对 \tilde{R}_{new} 进行识别诊断, 并输出诊断结果;
- 5) 返回 1), 重复上述步骤对下一待诊断的故障样本进行测试。

3 试验结果与分析

为了验证本文所提出的 Ce-ATRP-CNN 模型在实际轴承数据中的故障诊断性能, 采用美国凯斯西储大学 (CWRU, case western reserve university)^[20] 轴承数据集和华中科技大学 (HUST, Huazhong university of science technology) 轴承数据集进行试验验证。试验环境: 硬件环境为 12th Gen Intel (R) Core (TM) i5-12500H CPU; 软件环境为 Windows 11 系统, MATLAB R2023b 和 Python 3.9。

3.1 试验一 CWRU 轴承数据集

3.1.1 CWRU 数据描述

CWRU 数据集采自于由一个 2 马力的电动机、一个扭矩传感器/编码器、一个功率测试计和电子控制器组成的滚动轴承实验装置。其中, 轴承型号为 SKF-6205。

本研究选用采样频率为 12 kHz, 负载为 0 kW, 转速为 1 797 r/min 条件下轴承驱动端的一维时间序列数据作为试验对象。故障点分别在轴承的滚动体、内圈和外圈 3 个部位通过电火花加工出单点损伤直径为 0.007、0.014、0.021 inch 的故障, 共 10 种状态, 包

含 1 种正常状态和 9 种不同的故障状态, 具体故障类型、故障尺寸及标签设置如表 1 所示。

表 1 滚动轴承故障类型

| 故障类型 | 故障尺寸/inch | 符号 | 标签 |
|-------|-----------|-----|----|
| 正常状态 | — | H | 0 |
| 滚动体故障 | 0.007 | BF1 | 1 |
| | 0.014 | BF2 | 2 |
| | 0.021 | BF3 | 3 |
| 内圈故障 | 0.007 | IF1 | 4 |
| | 0.014 | IF2 | 5 |
| | 0.021 | IF3 | 6 |
| 外圈故障 | 0.007 | OF1 | 7 |
| | 0.014 | OF2 | 8 |
| | 0.021 | OF3 | 9 |

通过对 10 种状态分别进行分割采样, 使每个样本长度为 1 024, 每种状态均生成 200 个样本, 共 2 000 个样本。使用本文所述的 Ce-ATRP 方法对一维时间序列编码共得到 2 000 张二维故障特征图像。实验前按照训练集和测试集 7:3 的比例将图像随机分配, 每种故障状态训练集和测试集的分配数量分别为 140 和 60, 共 1 400 个训练集, 600 个测试集。

3.1.2 试验参数设置

CNN 网络^[21] 结构参数设置如表 2 所示, 其中, 各层激活函数均采用 ReLU 函数^[22], 参数优化器采用 Adam 自适应优化算法, 学习率设为 0.001, 采用交叉熵函数作为损失函数, 批量大小设置为 32, 迭代次数设置为 100。

表 2 CNN 网络参数设置

| 序号 | 网络层 | 参数属性 |
|----|-----------|----------------------|
| 1 | Input | 输入特征图尺寸 128×128×3 |
| 2 | Conv1 | 卷积核参数 5×5×32, 步长为 1 |
| 3 | Max_pool1 | 2×2 池化核, 步长为 2 |
| 4 | Conv2 | 卷积核参数 3×3×64, 步长为 1 |
| 5 | Max_pool2 | 2×2 池化核, 步长为 2 |
| 6 | Conv3 | 卷积核参数 3×3×128, 步长为 1 |
| 7 | Max_pool3 | 2×2 池化核, 步长为 2 |
| 8 | Conv4 | 卷积核参数 3×3×256, 步长为 1 |
| 9 | Fc1 | 2 560 个节点, 激活函数 ReLU |
| 10 | Fc2 | 10 个节点, 激活函数 Softmax |
| 11 | Output | 10 个节点, 对应 10 种类型 |

在训练过程中, 将训练集图像输入模型获取预测输出, 使用交叉熵损失函数计算预测值与真实标签之间的损失:

$$L = - \sum_{i=1}^c y_i \log(\hat{y}_i) \quad (13)$$

其中: C 为类别数; y_i 为真实标签向量中第 i 个元素

(若样本属于第 i 类, 则 $y_i = 1$, 否则 $y_i = 0$); \hat{y}_i 为模型预测的概率分布中第 i 个元素。

同时, 为了优化模型性能, 采用反向传播算法自动计算梯度, 并通过 Adam 优化器更新网络中的权重和偏置项, 通过最小化交叉熵损失, 使模型在 100 个训练周期中不断迭代和调整参数, 逐步提升分类性能。

每次迭代结束后, 在测试集上评估模型获取预测类别, 通过统计预测正确的样本数与总测试样本的比值计算故障诊断准确率。

3.1.3 试验结果与分析

α 值的选择过大容易造成故障特征被自适应阈值滤除, 过小则易造成噪声幅值的残余影响故障特征纹理从而会影响后续故障诊断准确率。为了确定 α 的最优参数值, 本文分别选取 $\alpha = 86、89、92、95、98$ 进行实验, 其训练过程中的测试准确率曲线如图 2 所示。由图可知, 当 $\alpha = 95$ 时, 训练准确率最高, 故本文选取 $\alpha = 95$ 作为最优百分位参数取值。

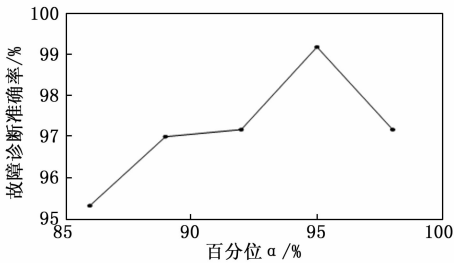


图 2 不同 α 值的故障诊断准确率

为了展示 ATRP 对不同工况下故障特征图像的可区分性, 本文给出了故障尺寸为 0.007 inch 的滚动体故障、内圈故障、外圈故障以及正常工况的 ATRP, 如图 3 所示。

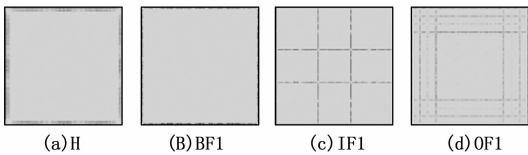


图 3 Ce-ATRP 图 ($\alpha = 95$)

由图 3 可知, 当 $\alpha = 95$ 时采用本文方法的图像编码方式转换的故障特征图像更直观地展示滚动轴承在不同故障状态下的振动特性, 每张图分别体现了四种故障类别不同的纹理特征, 使得不同类型的故障特征更加突出和易于识别。

训练集和测试集的准确率曲线和损失值曲线如图 4 所示。

由图 4 可知, 随着迭代次数的增加, 训练集和测试集的准确率不断上升并逐渐收敛, 经过 5 次迭代后分类准确率已经达到 80% 以上; 训练 45 次后准确率和损失

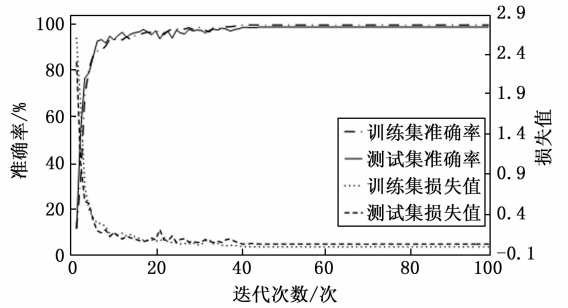


图 4 准确率和损失值曲线

曲线均趋于平稳, 最终滚动轴承故障诊断准确率达到 99.17%。为避免实验的偶然性, 采用相同方法和参数重复 5 次实验, 平均诊断准确率为 99%, 说明所提模型具有良好的稳定性。

测试集的混淆矩阵如图 5 所示。

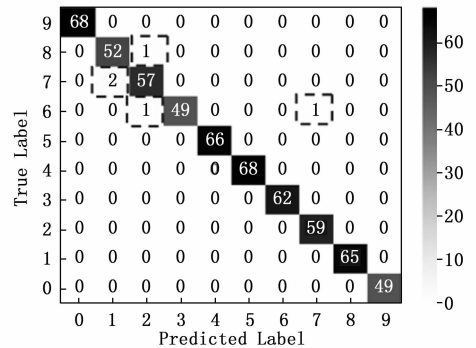


图 5 本文方法的混淆矩阵

标签 1、2、3 的故障状态出现误分类: 1 个标签 1 的样本被错误识别为标签 2, 2 个标签 2 的样本被错误识别为标签 1, 两个标签 3 的样本分别被错误识别为标签 2 和标签 7。具体分析可知: 除滚动体故障外, 其余类别的故障诊断准确率均为 100%。在 5 个错分类样本中, 仅 1 个样本出现故障类型识别错误 (滚动体故障被诊断为外圈故障), 其余 4 个样本则是故障程度识别有误, 但均诊断为滚动体故障, 这进一步验证了本文方法的有效性。

3.1.4 对比实验及分析

为了验证本文方法对于滚动轴承故障诊断的性能, 选取以下 3 组实验进行对比: 1) Ce-RP-CNN: 使用绝对距离对两个时间序列点的幅值直接编码转换为特征图像; 2) Ce-相空间 RP-CNN: 通过多次实验选择识别效果最好的参数 (嵌入维数为 3、延迟时间为 2) 进行相空间重构; 3) Ce-MTF-CNN。3 组实验均采用与本文相同的倒谱分析对原始信号进行预处理, 分别使用 RP、相空间 RP 和 MTF 转换方法将预处理后的一维振动信号转换为二维图片并输入与本文模型相同的网络训练和测试。采用相同的参数分别进行 5 次实验计算平均值,

结果如表 3 所示。

表 3 不同模型对比试验结果

| 对比试验 | 平均准确率/% |
|---------------|---------|
| Ce-ATRP-CNN | 99.00 |
| Ce-相空间 RP-CNN | 96.11 |
| Ce-RP-CNN | 95.50 |
| Ce-MTF-CNN | 95.28 |

由表 3 可知, 本文模型的平均诊断准确率最高, 与其他 3 种模型在准确率上相比分别提高了 2.89%、3.50%和 3.72%, 对比其他模型具有更高的诊断精度, 更适用于滚动轴承故障识别。

采用 3 种对比方法其中一次实验的诊断混淆矩阵如图 6 所示。

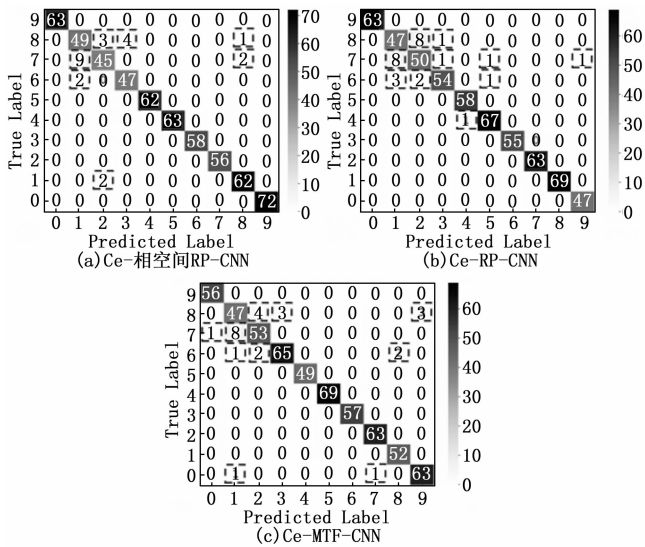


图 6 不同诊断模型的混淆矩阵

由图 6 可知, 3 种对比方法均有较多样本在滚动体部件发生不同故障程度间的误分类现象。除此之外, Ce-相空间 RP-CNN 模型和 Ce-MTF-CNN 模型分别有 5 个和 6 个样本出现滚动体和外圈的故障类型间的误分类情况; Ce-传统 RP-CNN 诊断模型有 3 个滚动体故障的样本被误分类为内圈故障和外圈故障。综上, 3 种对比模型相较于本文模型均有较多样本出现误分类, 这进一步验证本文模型具有良好的故障定位能力和故障程度识别能力。

为了更加清晰、直观地对本文模型的特征提取效果进行分析, 采用 t-SNE 算法对上述不同诊断模型进行特征可视化分析。如图 7 所示, 图中不同形状的点代表不同类别的故障。

由不同模型的特征可视化图可以直观地看出: 其他 3 种模型均有较为明显的混叠现象, 用肉眼很难区分出不同故障类别的聚类程度; 而本文所提模型的可视化结

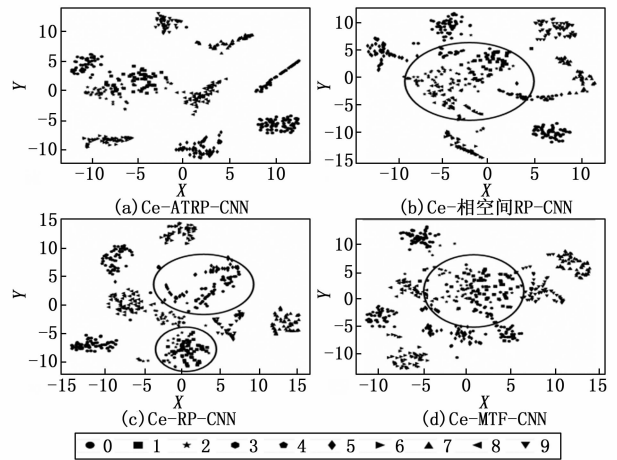


图 7 不同方法可视化结果图

果中显示出较为清晰的边界和可分性, 表明采用本文模型能够较为有效区分不同故障类别。

为了定量分析不同模型的特征可视化类间聚集程度, 选用 Fisher 评分作为量化指标, Fisher 评分通过类间散度矩阵的迹和类内散度矩阵的迹的比值计算, 其中, 类间距离反映类间分散程度, 类内距离反映的是类内紧凑程度。Fisher 评分越高说明故障类的分布差异越显著 (类间差异大), 且同一类别内部样本分布越紧密 (类内距离小), Fisher 评分具体公式表示为:

$$F = \frac{tr(S_b)}{tr(S_w)} \quad (14)$$

$$S_b = \sum_{i=1}^C \frac{n_i}{n} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (15)$$

$$S_w = \frac{1}{n} \sum_{i=1}^C \sum_{x_i \in C_i} (x_i - \bar{x}_i)(x_i - \bar{x}_i)^T \quad (16)$$

其中, S_b 为类间散度矩阵; S_w 为类内散度矩阵; $tr(\cdot)$ 为矩阵的迹; C 为类别数; n_i 表示第 i 类的样本数目; n 为样本总数; \bar{x}_i 表示第 i 类样本的中心点; \bar{x} 为所有样本中心点; C_i 表示第 i 类样本的集合; x_i 表示第 i 类样本点; T 为转置。

由表 4 可知: 本文方法的 F 值在所有方法中最高; 其中, 类间散度高于其他方法, 说明不同故障类间分布差异最明显; 并且类内散度最低, 说明在同一故障类中的所有样本分布紧密。这也验证了使用本文方法进行故障识别能够提升不同故障类的类间区分度, 同时也能够使类内样本更加聚合。

表 4 不同方法的聚类指标

| 不同方法 | $tr(S_b)$ | $tr(S_w)$ | F |
|---------------|-----------|-----------|-------|
| Ce-ATRP-CNN | 56 296.00 | 2 298.28 | 24.50 |
| Ce-相空间 RP-CNN | 50 957.44 | 7 310.52 | 6.97 |
| Ce-RP-CNN | 54 277.82 | 4 679.62 | 11.60 |
| Ce-MTF-CNN | 46 972.81 | 4 605.44 | 10.20 |

3.1.5 抗噪性试验

在实际应用中, 滚动轴承通常处于复杂的运行环境中, 为了验证本文模型在噪声条件下的适应能力, 对原始振动信号加入不同信噪比 (SNR, signal-to-noise) 的高斯白噪声, SNR 表示有效信号的功率与噪声功率的比值, 其公式表示为:

$$SNR(dB) = 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \quad (17)$$

其中: P_{signal} 为信号功率; P_{noise} 为噪声功率。

在原始信号上分别加入信噪比为 10、6 和 2 dB 的高斯白噪声, 采用与 3.1.4 节相同的参数进行 3 次实验取平均值, 结果如图 8 所示。

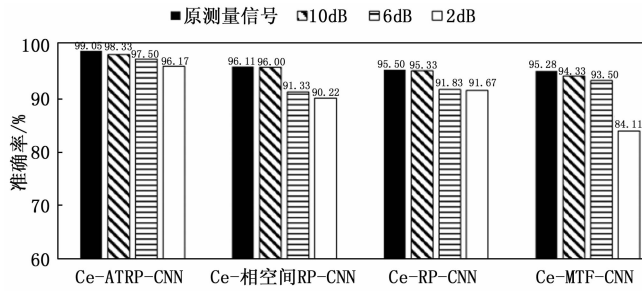


图 8 不同强度噪声下结果

如图 8 可知, 随着加入噪声的强度增大, 所有模型的故障诊断准确率均出现不同程度的下降, 但在 3 种不同噪声环境下 Ce-ATRP-CNN 的诊断准确率下降幅度较小并且始终在所有模型中的诊断准确率中保持最高。其他三种对比方法对于噪声较为敏感, 当加入信噪比为 6 dB 的高斯白噪声时, 诊断准确率相较于原测量信号的准确率均出现较大幅度的下降。当信噪比达到 2 dB 时, 其余 3 种模型的平均诊断准确率分别为 90.22%、91.72% 和 84.11%, 而本文所提模型的故障诊断准确率仍有 96.17%。实验表明: 相较于其他 3 种方法, 本文模型对于噪声具有良好的适应能力。

3.2 试验二 HUST 轴承数据集

为了验证 Ce-ATRP-CNN 模型的泛化性, 本节选取 HUST 轴承数据集进行实验评估, HUST 数据集是华中科技大学于 2024 年公开的轴承故障数据集, 数据采集于 ER-16K 深沟球轴承。该数据集包含正常、内圈中度故障、内圈重度故障、外圈中度故障、外圈重度故障、滚动体中度故障和滚动体重度故障, 共 7 种类型, 分别对应标签 0~6。

采用 3 900 rpm 转速工况条件进行实验, 使用本文方法生成 1 400 张故障特征图像, 按照 7:3 的比例随机划分为训练集和测试集: 980 张图片用于训练, 420 张图片用于测试。准确率曲线和损失曲线如图 9 所示。通过诊断模型进行训练和测试最终测试准确率为 99.05%。

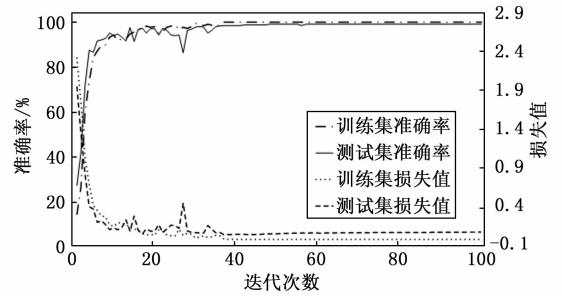


图 9 准确率和损失曲线

为了验证本文方法的泛化能力, 同样将 3.1.4 节的 3 种对比方法应用到 HUST 数据集中, 其故障诊断准确率如图 10 所示。由图可知, 本文模型相较于其他 3 种模型仍然取得了较好的诊断效果, 与其他 3 种方法在准确率上相比, 所提方法分别提高了 3.55%、1.91% 和 6.43%。

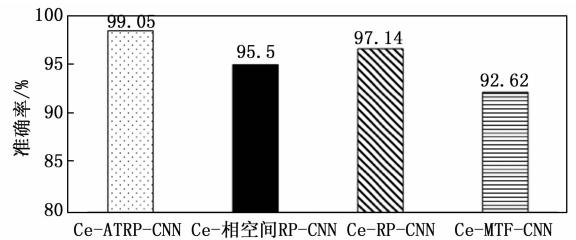


图 10 不同模型的准确率

上述不同方法的各类型故障诊断准确率如表 5 所示。

表 5 不同方法的各类型故障诊断准确率 %

| 数据类型 | Ce-ATRP-CNN | Ce-相空间RP-CNN | Ce-RP-CNN | Ce-MTF-CNN |
|-------|-------------|--------------|-----------|------------|
| 类型 0 | 100 | 87.9 | 94.7 | 87.9 |
| 类型 1 | 100 | 100 | 100 | 96.7 |
| 类型 2 | 98.4 | 93.8 | 98.5 | 93 |
| 类型 3 | 98.2 | 100 | 100 | 100 |
| 类型 4 | 100 | 94.3 | 93.1 | 86 |
| 类型 5 | 98.3 | 100 | 98.4 | 95.4 |
| 类型 6 | 98.1 | 88.7 | 96.4 | 89.3 |
| 平均准确率 | 99 | 94.9 | 97.3 | 92.6 |

本文模型应用于 HUST 轴承数据集的每种类型准确率均高于 98%, 对于七分类故障的平均诊断准确率为 99%, 结果表明: 应用不同的轴承数据集, 本文模型都能取得良好的诊断效果, 对比其他方法在故障识别效果上具有显著优势, 再次证明该方法对于滚动轴承故障诊断的有效性和可靠性。

4 结束语

本文针对滚动轴承故障振动信号在实际运行环境中故障特征易被噪声所掩盖的问题, 提出了基于 Ce-ATRP-CNN 的滚动轴承故障诊断方法: 通过倒谱分析对信

号频谱中较低幅值的分量给以较高的加权,从而有利于检测出弱周期信号,有助于从被噪声掩盖的故障振动信号中提取出故障特征。提出 ATRP 图像编码方法,将一维故障振动信号转换为二维故障特征图像,通过引入自适应阈值对振动信号状态点的距离矩阵进行筛选使故障特征更加突出和易于识别。为了验证本文模型对滚动轴承故障诊断的有效性和泛化性,分别采用西储大学和 HUST 轴承数据集进行验证,故障识别准确率分别达到了 99.17% 和 99.05%,使用两个不同的数据集均得到较高的故障识别准确率。同时,与不同方法进行对比可知,该方法对于滚动轴承具有较为优秀的故障识别和定位能力。虽然本文所提出的方法能够在不同轴承数据集中均取得良好的故障诊断效果,但仍然存在问题:本文仅对原测量信号加入弱噪声的条件进行验证,而未考虑对于强噪声环境下,即故障振动信号完全被噪声所淹没的情况。虽然本文方法相较于其他方法减少了多个参数选择的问题,但本文模型在确定自适应阈值时仍需手动设置百分位,如何消除百分位的选择对模型诊断的影响。

参考文献:

- [1] 段秉红,李勇,贺建军,等. 基于振动信号的轴承状态诊断方法研究 [J]. 化工机械, 2024, 51 (5): 801-806.
- [2] YAO G, WANG Y C, MOHAMED B, et al. A hybrid gearbox fault diagnosis method based on GWO-VMD and DE-KELM [J]. Applied Sciences, 2021, 11 (11): 4996.
- [3] 金江涛,许子非,李春,等. 基于 VMD 能量熵与优化支持向量机的轴承故障诊断 [J]. 计量学报, 2021, 42 (7): 898-905.
- [4] 刘磊,李舜酩,陆建涛. 基于卷积神经网络的旋转传动部件故障诊断综述 [J]. 机械设计, 2022, 39(10): 1-8.
- [5] 王运生,王黎明. 基于 SDP 图像和深度卷积网络的发动机故障诊断 [J]. 噪声与振动控制, 2023, 43 (5): 175-180.
- [6] LIU D D, CHENG W D, WEN W G. Rolling bearing fault diagnosis via STFT and improved instantaneous frequency estimation method [J]. Procedia Manufacturing, 2020, 49: 166-172.
- [7] 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, 2018, 65 (7): 5990-5998.
- [8] 杨腾,宁芊,陈炳才. 基于小波变换和残差神经网络模型的轴承故障诊断 [J]. 现代计算机, 2021 (15): 82-88.
- [9] 李恒,张氢,秦仙蓉,等. 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): 124-131.
- [10] 雷春丽,夏奔锋,薛林林,等. 基于 MTF-CNN 的滚动轴承故障诊断方法 [J]. 振动与冲击, 2022, 41 (9): 151-158.
- [11] 施保华,吴婷,赵子睿. 基于递归图和增强残差网络的轴承故障诊断 [J]. 轴承, 2024 (12): 87-94.
- [12] 万周,何俊增,姜东,等. 基于参数优化 SDP 分析的转子故障诊断方法 [J]. 振动与冲击, 2023, 42 (1): 81-88.
- [13] 范佳鹏,陈曦晖,李勇,等. 基于 MTF 和 AM-MSCNN 的滚动轴承故障诊断方法 [J]. 轴承, 2024, (12): 74-79.
- [14] 张龙,胡燕青,赵丽娟,等. 采用递归图编码技术与残差网络的滚动轴承故障诊断 [J]. 西安交通大学学报, 2023, 57 (2): 110-120.
- [15] SUN Y J, LI S H, WANG X H. Bearing fault diagnosis based on EMD and improved Chebyshev distance in SDP image [J]. Measurement, 2021, 176 (17): 109100.
- [16] 张天缘,孙虎儿,朱继扬,等. 基于 SDP 和改进 SAM-MobileNetv2 的滚动轴承故障诊断方法研究 [J]. 机械强度, 2024, 46 (4): 787-794.
- [17] DANIELLE D, ULISSES D, NATHALIA M, et al. Pixelwise remote sensing image classification based on recurrence plot deep features [C] //2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama: IGARSS, 2019.
- [18] RANDALL R B. A history of cepstrum analysis and its application to mechanical problems [J]. Mechanical Systems and Signal Processing, 2017, 97: 3-19.
- [19] WANG H, LIU Z L, PENG D D, et al. Attention-guided joint learning CNN with noise robustness for bearing fault diagnosis and vibration signal denoising [J]. ISA Transactions, 2022, 128: 470-484.
- [20] 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 (2): 100-131.
- [21] 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.
- [22] 田彪,张周锁,李想. 一种基于改进卷积神经网络的齿轮故障诊断方法 [J]. 机械设计, 2024, 41 (11): 1-7.