

梯度动态选择看齐的 DCFRC-FRNet 风电滚动轴承故障诊断

杨富龙^{1,2}, 巩丽俊¹, 吴 峰¹, 周津津¹, 杨 庆¹, 张 腾¹

(1. 兰州理工大学电气工程与信息工程学院, 兰州 730050; 2. 天水电气传动研究所集团有限公司, 天水 741020)

摘 要: 提出一种梯度动态选择看齐(DSA)、双通道特征融合技术(DCFRC)与深度学习相结合的风电滚动轴承故障诊断方法。首先,将采集到的一维时序数据采用递归图(RP)和马尔可夫转移场(MTF)转化为两种不同的二维图像数后据输入编码器重构为低维表示;然后,构建内嵌改进注意力机制 FRNet 的双通道特征融合卷积神经网络(DCFRC)提取深度特征信息,进行二级特征融合后完成特征分类;最后,利用解码器数据重构损失和特征提取网络分类损失构建梯度掩蔽矩阵,实现整个故障诊断过程梯度动态自适应。通过帕德伯恩数据集进行实验验证,结果表明该方法的诊断准确率为 99%,能有效提取故障特征信息。

关键词: 故障诊断; 滚动轴承; 风力发电机; 卷积神经网络; 梯度自适应

中图分类号: TH133.3; TP277

文献标志码: A

0 引 言

风力发电机的运行环境通常较恶劣,滚动轴承作为其中的关键组成部分,对整机的性能、稳定性和使用寿命具有显著影响。据统计,约有 50% 的风力发电机故障与滚动轴承有关。因此,及时准确地对风电机组轴承进行故障诊断,对于实现风电机组的预防性维护和状态检测具有重要意义^[1-2]。

深度学习作为机器学习的一个重要分支,将特征提取与故障分类整合为一体,实现端到端的架构模式,可直接从复杂的高维度数据中获取关键特征,成为故障诊断的流行方法之一^[3]。典型深度学习方法中的卷积神经网络(convolutional neural network, CNN)通过模拟人类视觉系统的机制,能从图像中自动学习和提取有用的特征,在处理图像处理方面的能力一直处于领先地位。因此,许多诊断方法都结合了数据预处理方法,将原始时域信号转换为二维图像后作为模型的输入。付文龙等^[4]利用连续小波变换将原始振动信号转化为二维时频图像,并结合数据增强技术与 CNN 实现故障诊断;王泽坤等^[5]针对传统深度学习模型存在长期依赖性、全局感知能力不足的问题,提出一种时频对称二维转换技术与多尺度 TransFusion 网络结合的故障诊断方法。以上方法均有效增强了模型性能。因此,为有效利用卷积神经网络的图像处理能力,本文将采用二维时频图作为特征提取模型的输入。

然而,基于深度学习的轴承故障诊断领域的大多数现有研究主要依赖于单通道二维时频图作为输入,在数据量较小

的情况下,单个域的故障信息有限,不足以提供全面的故障信息描述或实时了解设备运行情况。胡宝泉等^[6]针对单一域中包含特征信息有限的问题,将马尔可夫转移场、连续小波变化与双通道卷积神经网络相结合,实现了有限样本数据情况下的高精度故障诊断;姚德辰等^[7]在双通道特征融合网络的基础上结合动态卷积技术,在网络的更深层使用改进的 GhostNetV2 瓶颈层,通过动态卷积和注意力机制获得更多有益的特征。多域、多通道数据的有效利用大大有助于提高诊断模型的性能和精度,但多域数据对故障诊断模型的贡献良莠不齐,同时深层次复杂结构的网络还会带来梯度问题。

结合梯度自适应的多域、多通道信号融合成为推进轴承故障诊断的一条有前途的途径,其旨在更详尽地了解故障表现并进一步增强诊断效能。本文提出一种基于梯度选择看齐的 DSA-DCFRC-FRNet 的风电滚动轴承故障诊断方法。本研究的主要贡献体现在以下 3 个方面: 1) 基于递归图(recurrence plot, RP)和马尔可夫转移场(Markov transition field, MTF)在表达故障特征信息时的互补特性,构建双通道特征融合网络(dual channel feature fusion CNN, DCFRC)。其中 RP 通道通过相空间重构可有效捕获时间序列的周期性特征和突变点信息,而 MTF 通道则利用概率转移矩阵的纹理特征描述系统状态演化规律。2) 针对简单结构网络所提取特征表达能力欠缺的问题,创新性地提出改进型特征细化网络(feature refinement network, FRNet)。该网络在保留特征空间关系建模能力的同时,将传统注意力机制的计算复杂度大幅

收稿日期: 2025-01-17

基金项目: 甘肃省重大科技项目(25ZDWA003); 甘肃省联合科研基金(24JRRA829); 甘肃省重点研发计划-工业领域项目(25YFGA033)

通信作者: 杨富龙(1982—), 男, 博士、副教授、硕士研究生导师, 主要从事智能感知与信息处理等方面的研究。yangfulong1982@126.com

降低。3)针对多域数据对模型训练贡献不同及深度网络存在的梯度问题,提出梯度动态选择对齐算法。该算法通过联合优化自编码器的图像重构损失和分类网络的交叉熵损失,构建可微分掩蔽矩阵与梯度矩阵进行矩阵点积运算,实现反向传播过程中梯度流的自适应调控。

1 所提方法

1.1 梯度动态选择看齐

随着故障诊断模型的网络层数增加,故障信号的特征的提取能力也在不断增强。但在深层网络当中,由于链式法则的应用,梯度需通过多层进行反向传播,不合理的梯度值会影响整个模型的诊断性能。为更好地利用梯度来指导模型去关注和学习有益的样本信息,通过数据重构损失和分类网络的损失来构建梯度掩蔽矩阵,实现梯度动态选择看齐。

如图1所示为梯度动态选择看齐方法,整个过程包括二维原始信号重构、特征分类任务和梯度掩蔽,具体过程如下:

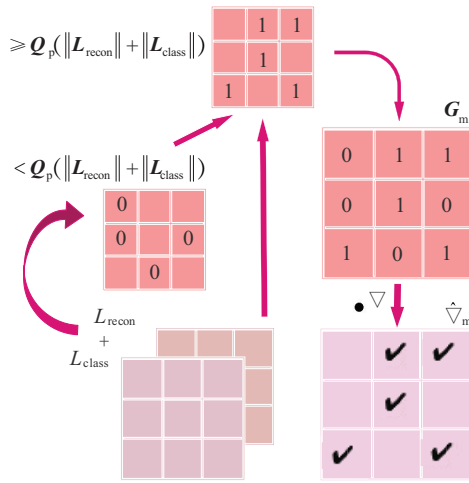


图1 梯度动态选择看齐算法

Fig. 1 Gradient dynamic selection of alignment algorithm

1)二维原始信号重构:将二维原始信号 a_i 输入编码器网络 E_c 压缩为低维表示 r_i ,之后解码器网络 D_c 通过转置卷积等上采样操作从提取的特征中恢复原始信号,如式(1)与式(2)所示。

$$r_i = E_c(a_i) \quad (1)$$

$$\hat{a}_i = D_c(r_i) \quad (2)$$

式中: i ——第 i 个二维图像数据; $\hat{a}_i$ ——解码器网络获得的重构信号。重构损失 L_d 用于比较解码器输出的重构信号与原始信号之间的区分度,以此展示自编码器特征提取及重构过程的质量,重构损失为:

$$L_{\text{recon}} = \sum_{i=1}^n \|\hat{a}_i - a_i\|^2 \quad (3)$$

2)特征分类任务:利用有标签的二维图像数据,通过故

障特征分类网络交叉熵损失来训练特征分类网络 F_c 预测的概率分布与真实标签概率分布之间的差异,对于一个真标签及其对应预测的概率分布。式(4)为分类网络的标签预测概率,式(5)为预测概率与真实标签的交叉熵损失。

$$\hat{b}_i = F_c(a_i) \quad (4)$$

$$L_{\text{class}} = -\sum_i^c b_i \ln(\hat{b}_i) \quad (5)$$

式中: b_i ——真实标签; $\hat{b}_i$ ——对应预测概率; c ——故障类别数量。

3)梯度掩蔽:为保证模型始终强调学习有益的样本,避免梯度问题,在各个训练步骤中,首先固定网络参数,之后通过更新掩蔽矩阵来选定适合的梯度进行反向传播以优化网络参数。梯度矩阵 ∇ 为重构损失及特征分类损失相对于编码器网络参数及特征分类网络参数的微分,而被掩蔽的梯度由梯度矩阵与掩蔽矩阵点积结果给出。式(6)~式(8)为掩蔽矩阵、梯度矩阵和梯度动态选择看齐矩阵的计算过程。

$$G_m = \begin{cases} 0, & \|L_{\text{recon}}\| + \|L_{\text{class}}\| \geq Q_p(\|L_{\text{recon}}\| + \|L_{\text{class}}\|) \\ 1, & \|L_{\text{recon}}\| + \|L_{\text{class}}\| < Q_p(\|L_{\text{recon}}\| + \|L_{\text{class}}\|) \end{cases} \quad (6)$$

$$\nabla_{\theta_{\text{enc}}} = \frac{1}{L_{\text{recon}}} \cdot \frac{\partial L_{\text{recon}}}{\partial \theta_{\text{enc}}}, \nabla_{\theta_{\text{class}}} = \frac{1}{L_{\text{class}}} \cdot \frac{\partial L_{\text{class}}}{\partial \theta_{\text{class}}} \quad (7)$$

$$\hat{\nabla}_m = G_m \cdot \nabla \quad (8)$$

DSA 动态选择看齐方法的详细过程如下:

DSA 方法

输入:重建损失 L_{recon} 和特征分类损失 L_{class}

输出:反向传播的选定梯度

1 for 全部批次中的一批次 do

2 L_{recon} 和 L_{class} 的归一化操作

3 计算归一化 L_{recon} 和 L_{class} 的总和

4 通过式(6)获得掩蔽矩阵 G_m

5 掩蔽矩阵 G_m 和梯度矩阵 ∇_{θ} 点积

6 获得梯度动态选择看齐矩阵 $\hat{\nabla}_m$

7 返回 $\hat{\nabla}_m$

深层网络中,反向传播时梯度需经过多层或时间步的连乘,此时若各层梯度长期偏离,则会导致其指数级衰减或增长。梯度动态选择看齐算法首先可根据不同层的重要性动态分配权重,平衡各层对总梯度的贡献,即允许梯度绕过部分层,从而减少连乘层数,直接传递到浅层;其次,通过调整梯度方向,使其与优化目标或主成分对齐,减少方向冲突导致的抵消效应,增强有效信号传递。梯度动态选择对齐算法通过动态干预梯度的方法和路径,有效解决了梯度爆炸与消失的问题,使得深层网络训练更加稳定。

1.2 递归图与马尔可夫转移场

1.2.1 递归图

递归图是 Eckmann 等基于庞加莱定理而建立起来的直观地反映非线性动态系统特性的一种方法^[8]。通过绘制式(9)中的递归矩阵,并用不同色深表示其二进制项即可获得递归图。

$$R_{ij}(\varepsilon) = \theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|), \quad i, j = 1, \dots, N \quad (9)$$

式中: N ——测量点 $\vec{x}_i$ 的数量; ε ——阈值距离; $\theta(\cdot)$ ——Heaviside 函数(如果 $x < 0, \theta(x) = 0$; 否则 $\theta(x) = 1$); $\|\cdot\|$ ——规范化操作。

$$\vec{x}_i \approx \vec{x}_j \Leftrightarrow R_{ij} \equiv 1 \quad (10)$$

如果 $R_{ij} \equiv 1$, 则在坐标 (i, j) 处绘制黑点, 如果 $R_{ij} \equiv 0$, 则绘制白点。递归图的两个坐标轴都是时间轴, 并显示向右和向上。此外, 根据定义, 递归图相对于主对角线是对称的, 即 $R_{ij} \equiv R_{ji}$ 。

1.2.2 马尔可夫转移场

首先, 将时间序列数据 $X = (x_1, x_2, \dots, x_T)$ 按其取值范围划分为 Q 个分位箱, 每个分位箱内的数据量相同。每个数据点 x_i 属于一个唯一的分位箱 q_i 。构建一个 $Q \times Q$ 的马尔可夫转移矩阵 W , 其中 W_{ij} 表示从分位箱 i 转移到分位箱 j 的频率。

$$W_{ij} = \frac{\sum_{\forall x \in q_i, y \in q_j, x+1=y} 1}{\sum_{j=1}^Q W_{ij}} \quad (11)$$

通过引入时间轴, 将矩阵 W 扩展到马尔可夫变换域 M 。数据集沿时间轴分为 Q 个区间, i 和 j 处的数据点属于相应的区间: q_i 和 q_j 。 M_{ij} 表示 $q_i \sim q_j$ 的转移概率。在获得马尔可夫转换矩阵 M 之后, 将矩阵中的元素作为像素, 从而完成马尔可夫转换的整个过程。其计算公式为:

$$M = \begin{bmatrix} w_{ij}|x_1 \in q_i, x_1 \in q_j & \cdots & w_{ij}|x_1 \in q_i, x_n \in q_j \\ w_{ij}|x_2 \in q_i, x_1 \in q_j & \cdots & w_{ij}|x_2 \in q_i, x_n \in q_j \\ \vdots & \ddots & \vdots \\ w_{ij}|x_2 \in q_i, x_1 \in q_j & \cdots & w_{ij}|x_2 \in q_i, x_n \in q_j \end{bmatrix} \quad (12)$$

1.3 特征细化网络

注意力需计算每个位置与其他所有位置之间的相似度, 为每个元素分配不同的重要性权重, 从而生成更精确的特征表示^[9]。但是当序列长度较长时, 需高昂的内存开销和计算成本。

针对上述问题, 提出改进注意力机制 FRNet, 通过多尺度卷积和归一化操作替代注意力机制中原有的线性连接及 Softmax 操作, 并加入轻量化的深度卷积。保留特征细化能力的同时大幅度降低传统自注意力的高计算复杂度, 使网络易于训练, 且输入输出的大小一致, 这意味着其可插入任何模型的特征提取部分。特征细化网络如图 2 所示, 具体过程如下:

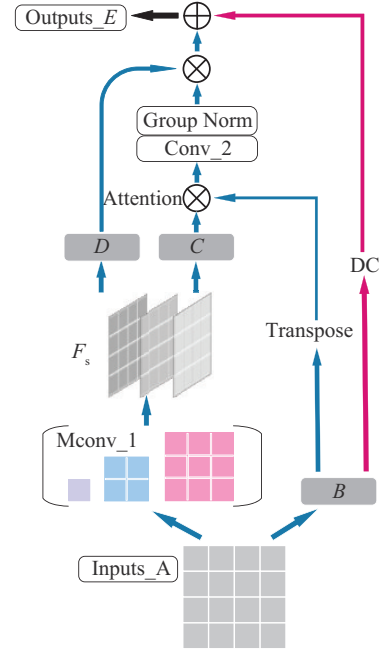


图2 特征细化网络 FRNet

Fig. 2 Feature refinement network FRNet

1) 首先将输入 $A_{Inputs} \in R^{H \times W}$ (为提取的图像数据) 送入多尺度卷积层捕捉不同尺度的特征, 随后将捕捉到的多尺度特征进行叠加整合得到 $F_s \in R^{H \times W}$, 使得模型可更好地理解特征的复杂结构。

2) 将输入 A_{Inputs} 和多尺度叠加特征 F_s 分别输入 1×1 卷积层和全连接层, 以获得特征 B 和特征 C , 其中 $B \in R^{H \times W}, C \in R^{H \times W}$ 。之后通过卷积和归一化得到注意力矩阵 $A_{tt} \in R^{H \times H}$, 如式(13)所示。

$$A_{tt,ij} = \frac{\exp(B_{xy} \cdot C_{ij})}{\sum_{i=1}^H \sum_{j=1}^W \exp(B_{xy} \cdot C_{ij})} \quad (13)$$

式中: $A_{tt,ij}$ ——特征 B 的 (x, y) 部分与特征 C 的 (i, j) 部分之间的关系。

3) 由注意力矩阵的计算公式可知, 其输出是权重与多尺度叠加特征 F_s 的加权和, 因此冗余的权重会导致特征的多样性下降, 进而导致模型的泛化能力下降。因此在原本的注意力矩阵中添加轻量化深度卷积 (deep convolution, DC) 来改善特征多样性。

将多尺度叠加特征 F_s 输入全连接层中得到特征 $D \in R^{H \times W}$, 由式(14)计算得到输出特征 $E_{Outputs}$ 。

$$E_{Outputs} = A_{tt} \cdot D + (A_{Inputs})_{DC} \quad (14)$$

改进后的注意力机制在减少计算资源消耗的同时, 对输入特征进行了多尺度的特征细化操作, 增强了特征的表达能力。其次, 通过生成自适应权重, 引导特征向更关键的方向学习, 增强了模型在故障分类任务中的性能表现。

1.4 双通道特征融合网络

在卷积神经网络的基础上使用双通道结合了 RP 可揭示

时间序列中的复杂结构的能力和 MTF 对状态转移的描述能力,同时对网络结构进行了深化,在每个通道末尾嵌入改进注意力机制来对提取的特征进行精细化加工,且其中的多通道融合可增强模型的自适应能力和鲁棒性。

双通道特征融合网络如图 3 所示,它以 RP 和 MTF 转化后的二维图像数据为输入,模型主要由两大部分组成:特征提取与融合部分和特征分类部分。首先通过两个 CNN-FRNet 通道对输入的数据信号进行特征提取,然后通过特征融合操作将提取的特征进行组合,其中特征融合有两个部分,

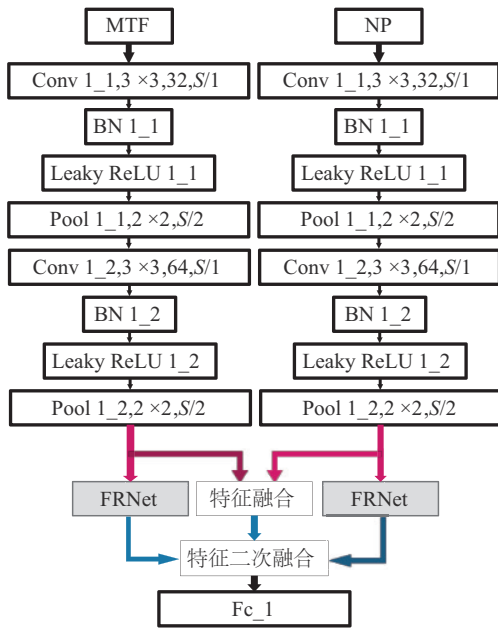


图 3 双通道特征融合网络
Fig. 3 Dual-channel feature fusion network

第一部分将两个 CNN-FRNet 通道中 CNN 所提取的数据信号特征进行融合,第二部分将两个 FRNet 注意力机制所加工过的数据序列特征与第一部分融合的特征再次融合得到最终的融合特征,最后通过全连接层及 Softmax 函数进行分类输出。

2 模型结构与诊断流程

2.1 诊断模型整体结构

为完全、细致的提取不同工作状态下的故障特征信息,从而完成对故障类别的准确判断,本文构建了动态选择看齐的 DSM-DCFFC-FRNet 风电滚动轴承故障诊断模型,如图 4 所示。

首先,为减少数据信息维度和提高故障特征的表征能力,将采集到的一维振动信号分别采用递归原理和马尔可夫转移场原理转化为二维图像数据,之后输入编码器将其压缩为低维表示,并使用 Resize 方法将压缩后的图片转化为 64×64 的小分辨率图片。其次,为提取二维图像数据中包含的深层次故障特征信息,利用所设计的双通道特征融合网络 DCFFC 提取输入样本特征,同时在每条通道中经过卷积操作后,引入特征细化网络 FRNet 对提取的特征进行最后的精细化加工。与此同时,利用解码器重构低维数据表示的损失和分类网络的特征分类损失来构建掩蔽矩阵,在与梯度矩阵点积后实现梯度动态选择看齐,即选取特定的梯度反向传播用来优化网络,以此使模型更加关注对整个训练过程有益的部分,并解决梯度爆炸或消失的问题。

最后,特征提取网络输出结果矩阵到全连接层之后通过 Softmax 层进行特征分类,输出最终的故障诊断结果。

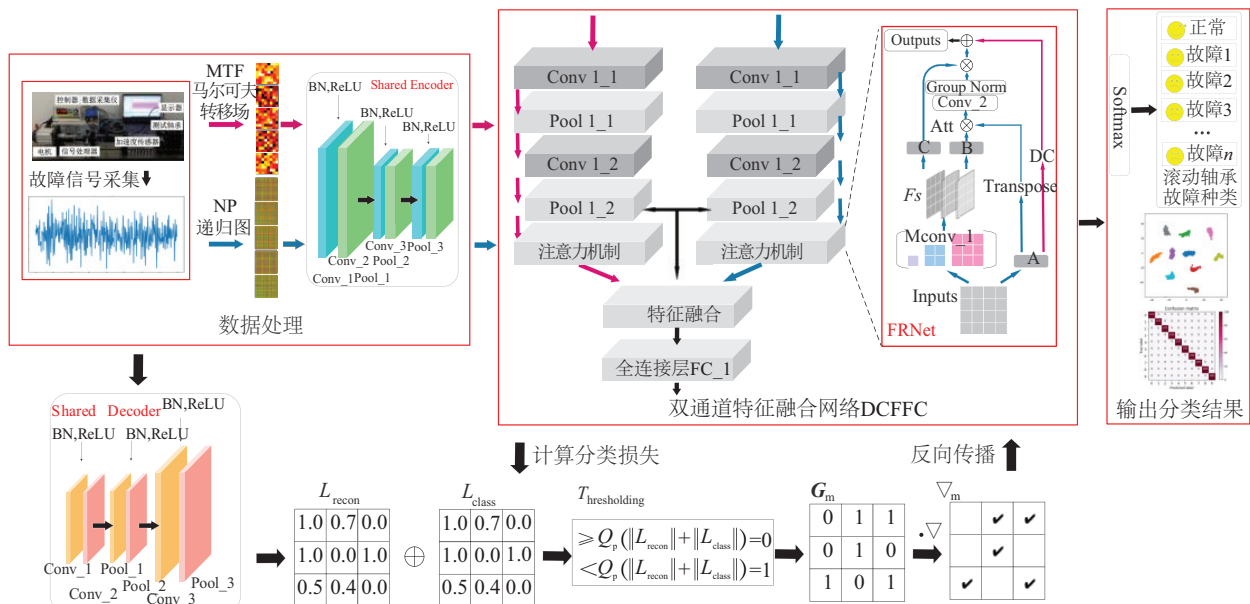


图 4 梯度动态选择看齐的 DSM-DCFFC-FRNet 模型结构

Fig. 4 DSM-DCFFC-FRNet model structure of gradient dynamic selection aligns

2.2 模型诊断流程

本文使用故障识别准确率和召回率作为故障诊断模型的量化指标,同时使用混淆矩阵和 t -SNE 可视化技术来直观展示故障标签与真实标签之间的接近程度和模型故障分类结果,提出的动态选择看齐的 DSM-DCFFC-FRNet 风电滚动轴承故障诊断模型流程如图 5 所示,具体步骤如下:

1) 进行风电滚动轴承故障信号采集,采用递归转化原理和马尔可夫转移场原理分别将采集到的故障信号转化为两组二维图像数据,之后利用编码器将其压缩为低维表示后通过 Resize 技术将其转化为 64×64 的图片,并将数据按 6:2:2 的比例划分为训练集、验证集和测试集。

2) 搭建 DCFFC-FRNet 特征提取网络,初始化模型的权重,并设置模型的梯度矩阵及其余超参数。

3) 利用所提的梯度动态选择看齐方法来反向传播选定的梯度,以此来优化模型结构和消除梯度爆炸或梯度消失的问题。

4) 使用验证集验证模型训练结果,检验梯度动态选择看齐方法的有效性。

5) 判断模型是否收敛,若收敛则进行下一步,否则返回步骤 2)。

6) 对训练好的故障分类模型 DSM-DCFFC-FRNet 进行测试,计算故障诊断的准确率和召回率并输出故障分类结果。

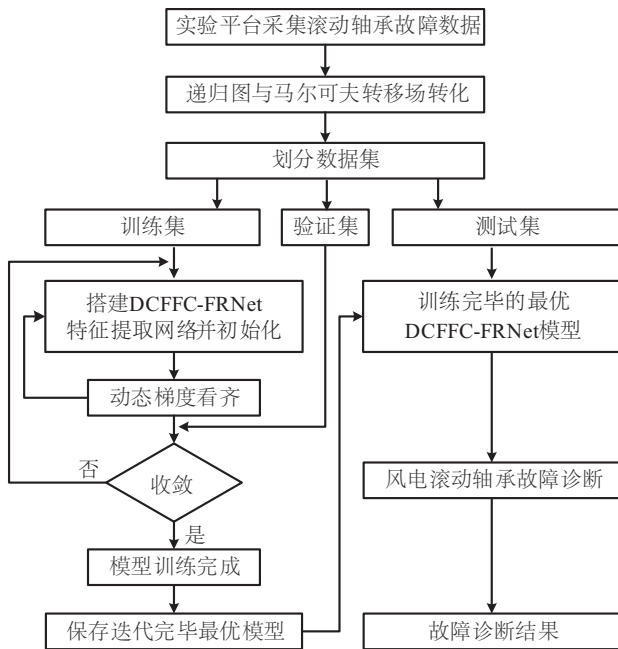


图5 模型诊断流程

Fig. 5 Model diagnosis flow

3 实验验证

3.1 帕德伯恩实验数据集

为验证本文所提故障诊断方法的有效性和优秀性能,德

国帕德伯恩(Paderborn)滚动轴承实验数据采集平台收集到的数据集,数据采集平台由电动机、扭矩测量轴、滚动轴承测试模块、飞轮和负载电机组成。通过将不同损伤类型的滚动轴承安装在轴承测试模块中生成实验数据,采样频率为 64 kHz。其中故障轴承分为人工损伤和真实损伤,故障位置为内圈和外圈。

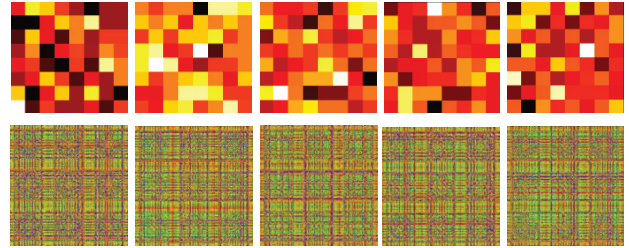


图6 状态信息的递归图和马尔可夫转移场图

Fig. 6 State information recurrence plot and Markov transition field diagram

利用 MTF 技术和 RP 技术将数据信息转化为二维特征彩色纹理图像,如图 6 所示,从左到右依次为人工槽沟、人工钻孔、人工点蚀、真实凹痕 4 种故障数据和正常运行数据转化后的二维图像。RP 技术可揭示时间序列中的复杂结构,获得定量的频率和时间信息,有助于确定故障发生的时间和频率;MTF 技术可通过颜色和纹理等元素来展示状态之间的转移概率,并准确描述信号变化的趋势。

数据集包含 4 种工况下的轴承故障数据,通过改变系统转速、负载扭矩及轴承的径向力来获取不同的工况条件。4 种工况的工作条件如表 1 所示,分别设置为工况 0~工况 3。

表1 工况操作参数

Table 1 Operating parameters for working conditions

工况编号	扭矩/(N·m)	转速/(r/min)	径向力/N
0	0.7	1500	1000
1	0.7	900	1000
2	0.1	1500	1000
3	0.7	1500	400

3.1.1 模型参数设置

在本章节中,使用帕德伯恩轴承数据集评估所提方法的故障诊断效果。实验的电脑配置为 AMD Ryzen 7-4600H 处理器,实验框架为 Keras-Python。

由图 4 可知,本文提出的 DSM-DCFFC-FRNet 包含 2 次特征融合,第一次特征融合是在两个通道的第二次池化操作完成后,将两个通道提取的特征进行融合。第二次特征融合是两个通道经过 FRNet 输出后,第一次融合后的特征与两个通道输出的经过特征细化后的特征数据进行融合。其中两个通道的输入是分别经过不同的二维图像转化技术转化后的彩色特征图像数据。在所有卷积操作后,使用零填充策略

保持特征映射大小不变。表 2 所示为 DCFFC-FRNet 特征提取网络的参数。

表 2 模型参数设置

Table 2 Model parameter settings

参数(Layer)	取值	填充	激活函数	输出大小
MTF 通道				
Input_1		same		64×64×1
Conv1_1	filters=32, k_s=3×3, str=1×1	same		64×64×32
Pool 1_1	k_s=2×2, str=2×2	same	ReLU	28×28×32
Conv1_2	filters=32, k_s=3×3, str=1×1	same	ReLU	28×28×32
Pool 1_2	k_s=2×2, str=2×2	same		14×14×32
FRNet				14×14×32
RP 通道				
Input_1				64×64×1
Conv1_1	filters=32, k_s=3×3, str=1×1	same	ReLU	64×64×32
Pool 1_1	k_s=2×2, str=2×2	same	ReLU	28×28×32
Conv1_2	filters=32, k_s=3×3, str=1×1	same		14×14×32
Pool 1_2	k_s=2×2, str=2×2	same		14×14×32
FRNet		same		14×14×32
特征融合层				
特征融合层 1				12544
特征融合层 2				25088
输出层				
全连接层	N_classes		Softmax	5

模型架构对于神经网络的复杂度与表示能力起着至关重要的作用。本研究采用系统性参数优化策略来提升诊断模型的性能表现。在参数选择阶段,首先基于网格搜索法对训练集进行五折交叉验证,通过权衡模型收敛速度与识别精度,初步筛选出 5 组候选参数组合。参数优化聚焦 3 个核心超参数:对比了批量大小分别为 8、16、32、64 和 128 时对模型诊断性能的影响;学习率在 0.01~1 范围内进行对数尺度采样,对比分析不同量级对故障诊断准确率迹的影响;卷积核尺寸选择 3×3、5×5、7×7、9×9 和 11×11 共 5 种规格,探究感受野扩展与特征提取能力的关联性,实验结果如图 7 所示。

实验设计严格遵循控制变量原则:当调整批量大小时,设置学习率为 0.05、卷积核尺寸为 7×7、迭代次数为 400;当考察学习率时,固定卷积核为 7×7、批量尺寸 32、迭代次数 400;当调整卷积核尺寸时,则保持学习率 0.05 及其他参数不变。每个参数组合均进行 3 次独立实验以消除随机性影响,最终通过综合评估验证集准确率指标,确定最优参数配置为

批量大小 32、学习率 0.05 和卷积核大小 7×7。

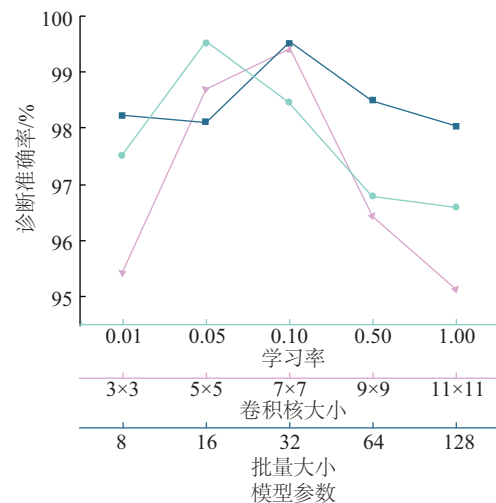


图 7 不同参数对模型诊断准确率的影响

Fig. 7 Influence of different parameters on accuracy of model diagnosis

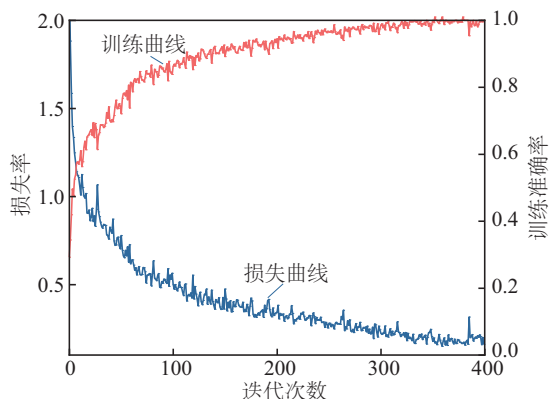


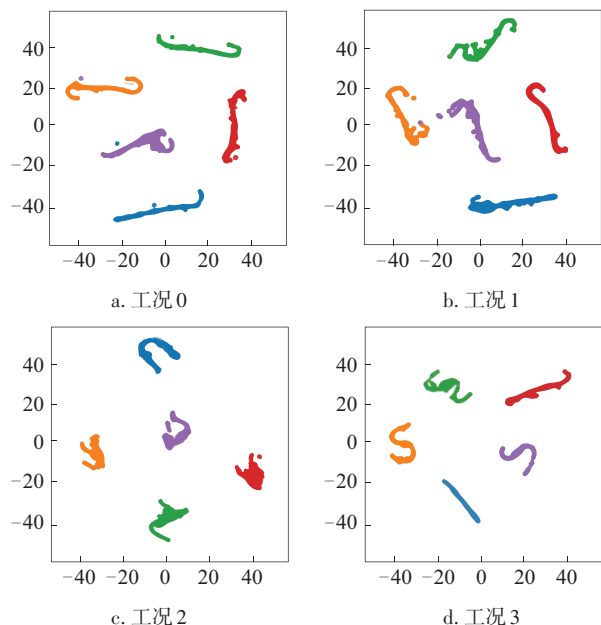
图8 模型训练迭代变化曲线

Fig. 8 Model training iteration change curves

此外,本研究借助训练集的准确率迭代曲线以及损失函数误差迭代曲线来验证所选取超参数的合理性。如图8所示,训练集的准确率迭代曲线呈现出基本稳定且持续保持在较高数值的状态,与此同时,损失函数误差迭代曲线也趋于稳定并维持在较低数值水平。这一结果充分表明,所选用的模型参数具有高度的合理性。

3.1.2 可视化实验结果与分析

模型的诊断效果会受到不同工况的影响,为进一步验证本文提出的 DSM-DCFFC-FRNet 模型的诊断能力及其泛化能力,因此分别对工况 0~工况 3 这 4 种工况下的故障数据进行实验,诊断结果采用 t -SNE 可视化技术直观展示,如图 9 所示, t -SNE 降维后二维坐标无实际物理意义,仅用于可视化聚类结构。

图9 工况0~工况3测试集的 t -SNE 可视化Fig. 9 t -SNE visualization for test set of case No.0-No.3

另外,引入混淆矩阵来详细分析每个类别故障信号的分类效果,如图 10 所示,水平轴代表预测输出分类,垂直轴代

表真实故障分类,对角线代表预测标签正确的准确率。

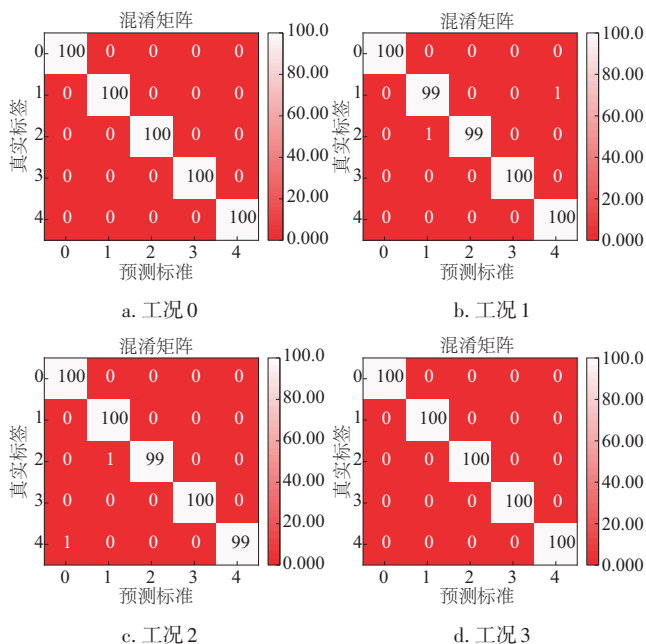


图10 工况0~工况3测试集的混淆矩阵

Fig. 10 Confusion matrix for test set of case No.0-No.3

综合上述对 4 中工况下的故障结果,4 中工况数据的诊断准确率的平均值都达到 100%,测试集的混淆矩阵表明除工况 1 和工况 2 下有 1% 的数据区分错误外,其余故障类型全部准确。该结果说明 DSM-DCFFC-FRNet 模型对于不同工况下的故障都具有很好的诊断效果,同时也说明模型具有良好的泛化性。

3.1.3 对比结果与分析

为验证特征细化网络 FRNet 和双通道特征融合在故障诊断方面的有效性,使用工况 0 的原始振动信号作为输入,与单通道模型和不使用特征细化网络 FRNet 的模型进行对比实验,得到的实验结果如表 3 所示。

表3 不同结构的故障诊断效果对比

Table 3 Comparison of fault diagnosis results of different structures

通道模式	是否嵌入 FRNet	二维图像 转化方式	准确率/ %	召回率/ %
单通道	是	MTF	92.36	92.03
	否		86.74	84.31
单通道	是	RP	94.44	93.99
	否		90.28	88.37
双通道	是	MTF、RP	98.77	98.40
	否		96.98	96.31
双通道	是	MTF、RP	100.00	100.00
	否		98.75	98.13

由表 3 可知,单通道模式下嵌入 FRNet 相比于未嵌入时故障诊断的准确率上升了 5.62 和 4.16 个百分点,双通道模式下嵌入 FRNet 相比于未嵌入时故障诊断的准确率上升了 1.79 和 1.25 个百分点。双通道模式相比于单通道模式故障诊断的准确率上升了 4.33~7.64 个百分点,双通道模式下使用二维图像转化相比于不使用二维图像转化故障诊断准确率上升了 1.23 和 1.77 个百分点。以上结果表明,本文设计并使用的二维图像转换、FRNet、双通道特征融合网络都对提升故障诊断性能有极大的正向作用。

为验证本文方法的故障诊断效果,在该数据集下,使用 LSTM、RP-CNN^[10]、MTF-GNN^[11]、P-2DCNN^[12]、DCNN-BiLSTM^[13]、DC-SACNN^[14] 共 6 网络与本文所提出的网络进行性能比较,对比结果如表 4 所示。

表 4 不同模型的对比实验结果

Table 4 Comparative experimental results of different models

模型名称	时间/ s	训练集		测试集	
		准确率/	召回率/	准确率/	召回率/
		%	%	%	%
LSTM	29	92.43	91.63	78.53	71.71
RP-CNN	46	94.86	94.41	87.21	82.72
MTF-GNN	137	96.35	94.64	95.55	92.68
P-2DCNN	76	95.03	93.54	92.19	91.81
DCNN-BiLSTM	109	97.42	93.27	94.60	93.41
DC-SACNN	49	98.30	97.46	97.51	96.85
本文方法	89	100.00	99.10	100.00	99.83

由表 4 可知,本文提出的方法故障诊断准确率最高,与其他 6 种方法相比分别提高了 7.57、5.14、3.65、4.97、2.58、1.70 个百分点。单分支简单结构的故障诊断模型诊断精度较低,原因是其只能提取局部特征,而二维输入信号方法诊断精度较低的原因是特征提取结构太简单,无法提取输入信号中的深层次特征信息。实验验证表明,本文提出的方法弥补了传统方法数据偏差敏感和难以提取非线性平稳信号特征的缺陷。

3.2 西安交通大学 Gearbox 数据集

轴承故障数据来自西安交通大学的 Gearbox 数据集,实验平台由交流机、转轴、控制器、测试轴承等组成。该数据集中包含的滚动轴承的故障状态分别为滚珠故障、内圈故障、外圈故障、混合故障、缺齿故障、断齿故障、齿根裂纹故障和齿面磨损故障共 8 种。

采样频率为 20480 Hz,正常状态及每种故障信号都取 500 样本,每个样本包含 1024 个采样点,并以 6:2:2 的比例划分训练集和测试集。数据集包含 3 种工况下的轴承故障数据,通过改变系统转速和轴承的径向力来获取不同的工况条件。3 种工况的工作条件如表 5 所示,分别设置为工况 0~工况 2。

利用递归图和马尔科夫转移场两种转换方法将原始振动信号转化为二维时频图像,得到 9 种数据二维图像如图 11 所示。从左到右、从上到下依次为 8 种不同故障数据和正常运行状态下的数据。

表 5 工况操作参数

Table 5 Operating parameters for working conditions

工况编号	转速/(r/min)	径向力/kN
0	2100	12
1	2250	11
2	2400	10

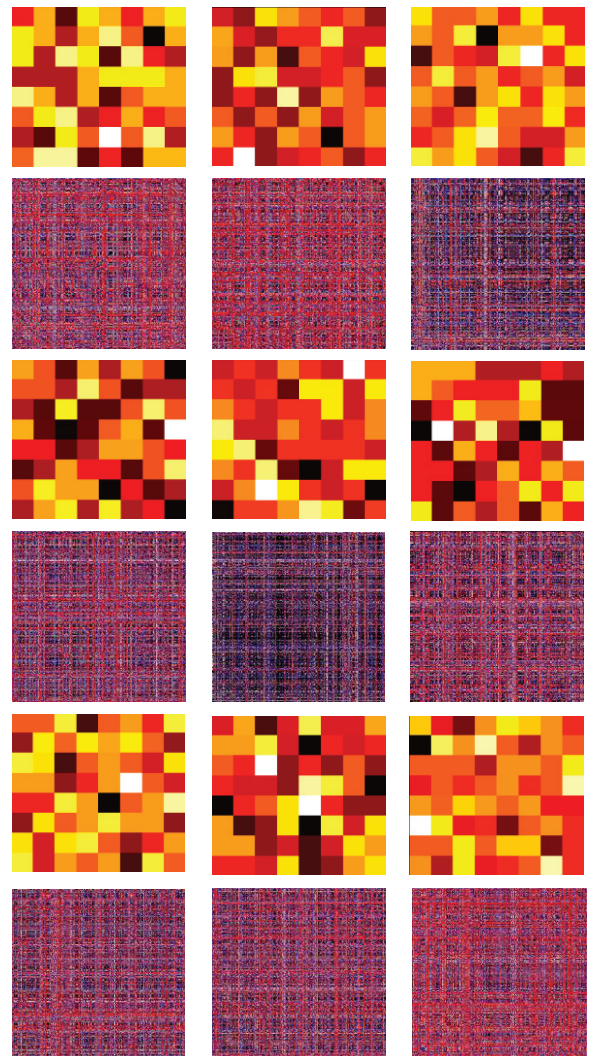


图 11 状态信息的递归图和马尔科夫转移场图

Fig. 11 State information recurrence plot and Markov transition field diagram

3.2.1 对比实验结果与分析

实验数据来自工况 0~工况 2 共 3 种工况环境下,实验设置固定批量大小为 32,经预实验验证可平衡显存占用与梯度稳定性,采用 Adam_W 优化器并设置基础学习率为 0.05。本节设置在西安交通大学 Gearbox 轴承数据集下与其他方法进

行对比,每组实验独立进行 10 次重复训练,每次训练包含 400 个完整迭代周期,平均准确率实验对比结果如图 12 所示。本文方法诊断准确率约为 98%,不同数据集中取得的优良故障结果证明 DSA-DCFFC-FRNet 故障诊断模型具备较强的泛化能力。

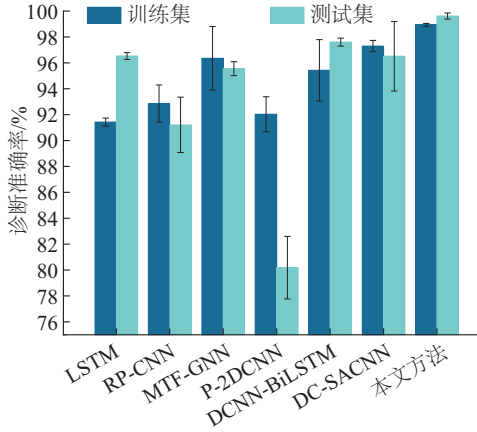


图 12 不同模型的对比实验结果

Fig. 12 Comparative experimental results of different models

3.2.2 可视化实验结果与分析

如图 13 所示的 t -SNE 可视化结果从特征空间层面揭示了模型的诊断机理,其中同类故障在不同工况下的特征向量余弦相似度均值达到 0.98,显著高于传统 CNN 模型,而不同故障类型间的特征距离保持在相对较大的稳定区间。

混淆矩阵定量解析了模型的细粒度分类能力:整体诊断准确率达到 99%,其中对角线元素(正确分类比例)均超过 95%,仅在少数几类形态相似故障间出现可解释的误判,其交叉错误控制在 0.11%~0.24% 范围内,如图 14 所示。

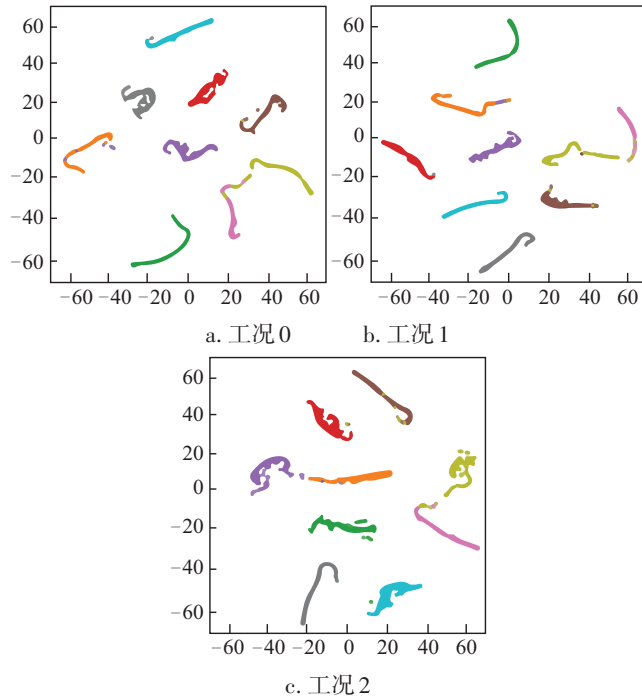
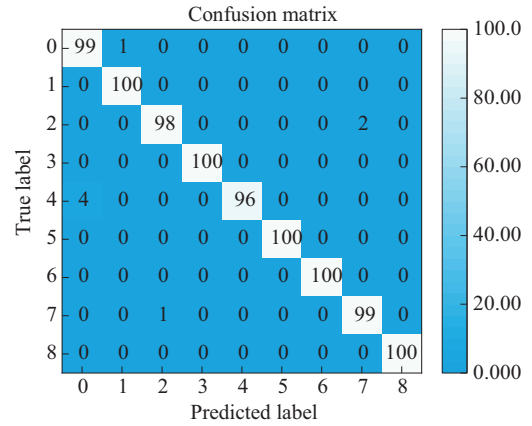
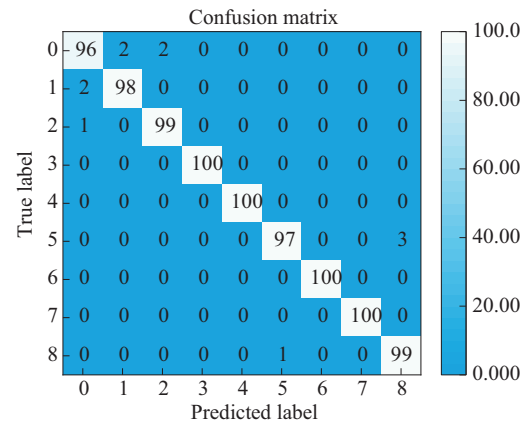


图 13 工况 0~工况 2 测试集的 t -SNE 可视化

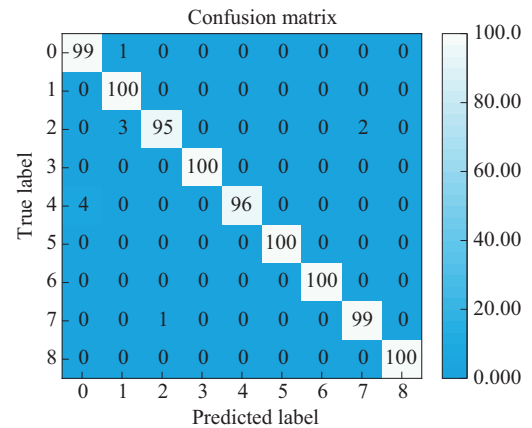
Fig. 13 t -SNE visualization for test set of case No.0-No.2



a. 工况 0



b. 工况 1



c. 工况 2

图 14 工况 0~工况 2 测试集的混淆矩阵

Fig. 14 Confusion matrix for test set of case No.0-No.2

4 结论与展望

针对复杂多变的工况下故障特征提取能力差的问题,结合了梯度动态选择看齐方法、双通道特征融合技术和深度学习,提出一种动态选择看齐的 DSM-DCFFC-FRNet 风电滚动轴承故障诊断方法,通过实验验证本文所提方法能有效实现风电滚动轴承的早期故障诊断。并得出以下主要结论:

1) 根据增强特征的表示能力和降低计算量的目标设计

了 FRNet, 相比于传统注意力机制, 在降低了计算量和计算复杂度的同时增强了故障特征的代表能力。

2) 双通道特征融合技术将 MTF 技术和 RP 技术转化后的二维图像数据中的深层次特征提取并加以融合, 巧妙结合了解决了多工况情况下滚动轴承振动信号难以有效全面提取和多样特征互补的问题。

3) 提出梯度动态选择看齐算法, 可在不同工况的背景下自适应选取特定梯度反向传播, 使模型始终关注有益的样本部分。同时也避免了复杂模型结构可能会产生的梯度爆炸或消失的问题。

尽管本文提出的 DSM-DCFFC-FRNet 方法在风电滚动轴承故障诊断中表现出较高的准确性和鲁棒性, 但仍存在以下局限性: 计算复杂度与实时性限制: 在处理长序列数据或高分辨率二维图像时, 多尺度卷积、注意力机制与梯度掩蔽操作可能导致训练和推理时间延长, 对实时性要求较高的工业场景存在挑战; 大规模数据适应性不足: 在面对更大规模或更高维度的工业数据时, 模型的计算效率与内存占用可能成为瓶颈。双通道并行处理与多级特征融合可能导致资源消耗呈非线性增长; 多模态数据融合缺失: 当前方法仅针对振动信号进行分析, 未整合声发射、温度或油液监测等多源异构数据, 限制了其对复杂故障模式的全面表征能力。

针对上述局限性, 后续研究可从以下方向展开: 轻量化模型设计: 优化 FRNet 与 DCFFC 的网络结构, 引入深度可分离卷积或知识蒸馏技术, 进一步降低计算复杂度; 探索动态剪枝策略, 在训练过程中自动剔除冗余参数; 分布式计算与硬件加速: 利用 GPU 并行计算或专用硬件加速双通道特征提取与梯度动态选择过程, 提升模型在大规模数据下的可扩展性; 多模态信息融合: 扩展模型输入维度, 融合声学、热力学等多模态信号, 并结合图神经网络构建跨模态关联, 以提升复杂工况下的故障诊断精度。

[参考文献]

- [1] 余萍, 宋紫琼, 曹洁, 等. 基于 BNN-RA 模型的风电机组轴承故障诊断研究[J]. 太阳能学报, 2025, 46(3): 643-651.
YU P, SONG Z Q, CAO J, et al. Research on fault diagnosis of wind turbine bearing based on BNN-RA model [J]. Acta energiae solaris sinica, 2025, 46(3): 643-651.
- [2] 邢作霞, 张玥, 郭珊珊, 等. 基于改进卷积神经网络的风电机组叶片覆冰诊断方法研究[J]. 太阳能学报, 2025, 46(3): 661-667.
XING Z X, ZHANG Y, GUO S S, et al. Diagnosis of blade icing based on improved convolution neural network in wind turbine study [J]. Acta energiae solaris sinica, 2025, 46(3): 661-667.
- [3] 邵海东, 肖一鸣, 颜深. 仿真数据驱动的改进无监督域适应轴承故障诊断[J]. 机械工程学报, 2023, 59(3): 76-85.
SHAO H D, XIAO Y M, YAN S. Simulation data-driven enhanced unsupervised domain adaptation for bearing fault diagnosis[J]. Journal of mechanical engineering, 2023, 59(3): 76-85.
- [4] FU W L, JIANG X H, LI B L, et al. Rolling bearing fault diagnosis based on 2D time-frequency images and data augmentation technique [J]. Measurement science and technology, 2023, 34(4): 045005.
- [5] WANG Z K, XU Z F, CAI C, et al. Rolling bearing fault diagnosis method using time-frequency information integration and multi-scale TransFusion network [J]. Knowledge-based systems, 2024, 284: 111344.
- [6] HU B Q, LIU J, ZHAO R Z, et al. A new dual-channel convolutional neural network and its application in rolling bearing fault diagnosis [J]. Measurement science and technology, 2024, 35(9): 096130.
- [7] YAO D C, ZHOU T, YANG J W, et al. Fault diagnosis of rolling bearings based on dynamic convolution and dual-channel feature fusion under variable working conditions [J]. Measurement science and technology, 2024, 35(6): 066110.
- [8] 成洁, 李思燃. 基于递归图和局部非负矩阵分解的轴承故障诊断[J]. 工矿自动化, 2017, 43(7): 81-85.
CHENG J, LI S R. Bearing fault diagnosis based on recurrence plots and local non-negative matrix factorization [J]. Industry and mine automation, 2017, 43(7): 81-85.
- [9] CUI L, TIAN X C, WEI Q Z, et al. A self-attention based contrastive learning method for bearing fault diagnosis [J]. Expert systems with applications, 2024, 238: 121645.
- [10] BAI R X, MENG Z, XU Q S, et al. Fractional Fourier and time domain recurrence plot fusion combining convolutional neural network for bearing fault diagnosis under variable working conditions [J]. Reliability engineering & system safety, 2023, 232: 109076.
- [11] WANG H T, LIU Z L, LI M J, et al. A gearbox fault diagnosis method based on graph neural networks and Markov transform fields [J]. IEEE sensors journal, 2024, 24(15): 25186-25196.
- [12] WANG J X, WANG D Z, WANG S H, et al. Fault diagnosis of bearings based on multi-sensor information fusion and 2D convolutional neural network [J]. IEEE access, 2021, 9: 23717-23725.
- [13] WANG Y, WANG Q R, ZHOU Y T. Improved Dual-channel CNN-BiLstm rolling bearing fault diagnosis study [C]//EEI 2022; 4th International Conference on

Electronic Engineering and Informatics. Guiyang, China, 2023: 1-5.

- [14] WANG D L, LI Y H, LU C, et al. Research on digital twin-assisted dual-channel parallel convolutional neural

network-transformer rolling bearing fault diagnosis method [J]. Proceedings of the Institution of Mechanical Engineers, Part B: journal of engineering manufacture, 2024: 09544054241290573.

GRADIENT DYNAMIC SELECTIVE ALIGNMENT FOR DCFFC-FRNET WIND TURBINE ROLLING BEARING FAULT DIAGNOSIS

Yang Fulong^{1,2}, Gong Lijun¹, Wu Feng¹, Zhou Jinjin¹, Yang Qing¹, Zhang Teng¹

(1. School of Electrical Engineering and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;

2. Tianshui Electric Drive Research Institute Co., Ltd., Tianshui 741020, China)

Abstract: A fault diagnosis method for wind turbine rolling bearings is proposed, which combines the gradient dynamic selective alignment (DSA) method, dual-channel feature fusion technique (DCFF), and deep learning. Firstly, the acquired one-dimensional time series data are transformed into two different two-dimensional image representations using recursive mapping (RP) and Markov transfer field (MTF). These transformed images are then input into an encoder for reconstruction into a low-dimensional representation. Next, a dual-channel feature fusion convolutional neural network (DCFFC) is constructed, embedding an improved attention mechanism within FRNet to extract in-depth feature information. A second-level feature fusion is performed to complete the feature classification. Finally, a gradient masking matrix is constructed using the data reconstruction loss of the decoder and the feature extraction network classification loss to achieve gradient dynamic self-adaptation throughout the fault diagnosis process. Experimental validation using the Paderborn dataset demonstrates that the proposed method achieves a diagnostic accuracy of 99%, effectively extracting fault-specific diagnostic information.

Keywords: fault diagnosis; rolling bearing; wind power generator; convolutional neural network; gradient adaptive