

基于 ResNet-RFA 和 BiGRU-SATT 的 弧齿锥齿轮箱故障诊断*

王晓晨¹ 许 昕² 潘宏侠² 张珂铭¹ 程 凯¹

(1. 中北大学机械工程学院 太原 030051; 2. 中北大学系统辨识与诊断技术研究所 太原 030051)

摘 要: 针对传统故障诊断模型在齿轮箱振动信号时空特征协同挖掘不足和噪声环境下鲁棒性较弱的问题, 本文提出一种融合感受野注意力残差网络(ResNet-RFA)与双向门控循环单元-自注意力机制(BiGRU-SATT)的智能故障诊断方法。首先对原始振动信号进行短时傅里叶变换(STFT)生成时频图像, 并保留一维时序信号以保留多维度信息; 随后构建双通道特征提取网络: 空间特征通道采用 ResNet-RFA 模块, 通过感受野注意力机制聚焦时频图像的关键区域; 时序特征通道采用 BiGRU-SATT 模块, 结合双向门控循环单元与自注意力机制捕捉动态依赖关系; 最后通过特征拼接融合策略整合时空信息, 输入全连接层实现故障分类。实验结果表明, 该方法准确率达到 100%, 显著优于对比模型(Transformer: 91%, Mamba: 96%, SVM: 94%, DBN: 89%), 在添加 10、20 dB 高斯-脉冲混合噪声后仍保持较高识别性能, 展现出优异的抗噪性与稳定性。综上, ResNet-RFA 与 BiGRU-SATT 的融合模型能够有效协同挖掘信号的时空特征, 在准确率和鲁棒性方面均优于其他对比模型, 适用于复杂工业环境。

关键词: 残差网络; 感受野注意力; 双向门控循环单元

中图分类号: TN06 **文献标识码:** A **国家标准学科分类代码:** 460

Fault diagnosis of spiral bevel gearbox based on ResNet-RFA and BiGRU-SATT

Wang Xiaochen¹ Xu Xin² Pan Hongxia² Zhang Keming¹ Cheng Kai¹

(1. School of Mechanical Engineering, North University of China, Taiyuan 030051, China;

2. System Identification and Diagnosis Technology Research Institute, North University of China, Taiyuan 030051, China)

Abstract: To address the limitations of traditional fault diagnosis models in collaboratively extracting spatiotemporal features from gearbox vibration signals and their weak robustness in noisy environments, this paper introduces an intelligent fault diagnosis method that combines a receptive field attention residual network (ResNet-RFA) and a bidirectional gated recurrent unit with self-attention (BiGRU-SATT). The process begins by converting raw vibration signals into time-frequency images via short-time fourier transform (STFT), while preserving the original 1D time-series data. A dual-channel network is then constructed: One channel uses ResNet-RFA to extract key spatial features from the time-frequency images, and the other uses BiGRU-SATT to capture temporal dependencies. The spatiotemporal features are merged and fed into a fully connected layer for classification. Experimental results demonstrate a high accuracy of 100%, outperforming comparison models (Transformer: 91%, Mamba: 96%, SVM: 94%, DBN: 89%) and showing strong noise robustness under 10 and 20 dB Gaussian-impulse mixed noise. In conclusion, the fusion model of ResNet-RFA and BiGRU-SATT can effectively and collaboratively mine the spatiotemporal features of signals, demonstrating superior accuracy and robustness over other comparative models, making it suitable for complex industrial environments.

Keywords: residual network; receptive field attention; bidirectional gated recurrent unit

0 引 言

弧齿锥齿轮箱作为机械设备中的关键部件, 其运行状

态直接关系到整个传动系统的可靠性与安全性。然而, 在实际运行过程中, 高速、重载等工况常诱发强噪声, 导致故障特征难以有效分离, 使诊断面临严峻挑战。因此, 突破强

收稿日期: 2025-10-21

* 基金项目: 内燃机可靠性国家重点实验室基金(skler-201911)项目资助

噪声下信号提取与识别的技术瓶颈显得尤为重要。

长期以来,国内外学者针对提高齿轮箱故障识别问题开展了深入的研究。在信号分解方法方面,苏缪诞^[1]对经验模态分解算法(empirical mode decomposition, EMD)进行改进,并将其应用于旋转机械故障诊断中,提升了非稳信号的处理能力;张建峰等^[2]提出基于加权双域稀疏分解模型的编码器信号特征提取方法,通过联合时域周期性组稀疏性与频域谐波稀疏性,实现了对微弱故障冲击与谐波干扰的高精度分离。在特征提取方面,别锋锋等^[3]则结合参数优化变分模态分解(variational mode decomposition, VMD)与核极限学习机(kernel extreme learning machine, KELM),增强了齿轮箱故障特征的提取效果。随着深度学习技术的发展,基于卷积神经网络(convolutional neural network, CNN)的方法备受研究者关注。Guo 等^[4]提出了一种基于 CNN 的旋转机械故障诊断新方法,验证了其在端到端诊断中的有效性。此外,深度置信网络(deep belief network, DBN)也被用于齿轮箱故障诊断^[5],拓展了深度模型的应用范围。在序列建模方面,基于神经网络的模型在时序信号建模中得到了改进。Wang 等^[6]提出了一种结合注意力机制的长短期记忆网络(long short-term memory, LSTM)模型,有效提升了故障特征的辨识能力;张雄等^[7]则利用门控循环单元(gated recurrent unit, GRU)与卷积网络结合,在复杂工况下实现了对时序依赖的精细刻画;苟小伟等^[8]将深度卷积神经网络(deep convolutional neural network, DCNN)与双向长短期记忆网络(bidirectional long short-term memory, Bi-LSTM)结合,进一步拓展了深度学习在时序信号中的应用。然而,深度学习也暴露出在训练过程中容易出现的退化问题,而残差网络^[9](residual network, ResNet)的引入则有效缓解了这一困境。魏秀业等^[10]基于特征融合与 ResNet 对齿轮箱进行故障诊断,验证了 ResNet 的有效性。为提升模型对关键特征的聚焦能力,韩延等^[11]将时序注意力引入原型网络,增强了模型对关键时间步的聚焦能力;王翰伟等^[12]提出基于注意力机制和 ResNet 的齿轮箱故障诊断方法,其引入显著提升了特征提取能力。与此同时,吴启航等^[13]将乘法与卷积网络相结合,通过设计可学习的滤波核提取故障频域模态,实现了高精度、低延迟的在线诊断。

综上,现有方法在齿轮箱故障诊断中取得了显著进展,但仍存在不足:多数模型在时空特征协同方面尚未充分挖掘,难以同时捕捉局部空间模式与长程时序依赖;强噪声下,模型的鲁棒性与泛化能力仍有待提升;现有方法缺乏对内部结构与时序关键节点的精细化注意力加权。

针对上述问题,本研究提出基于感受野注意力残差网络(receptive field attention residual network, ResNet-RFA)与双向门控循环单元-自注意力机制(bidirectional gated recurrent unit with self-attention, BiGRU-SATT)的智能故障诊断方法。通过双通道结构分别处理时频图像和

一维振动序列,结合 RFA 机制实现参数级感受野加权,增强空间特征的判别性;利用 BiGRU-SATT 捕捉双向时序依赖,并通过自注意力聚焦关键时间节点;最终通过特征拼接整合时空信息,提升在复杂噪声环境下的诊断精度与鲁棒性。

1 数据预处理

1.1 信号重叠采样

为解决深度模型训练中标签数据不足导致的过拟合风险,本研究引入了重叠采样扩容技术。其核心原理是利用数据样本内部存在的连续性与局部相关性。通过设计特定的重叠步长对单个样本进行多次重复采样,能从一个原始样本生成多个衍生样本,从而高效率地扩充训练集规模,最终提升模型的泛化能力,具体方法为图 1 所示。

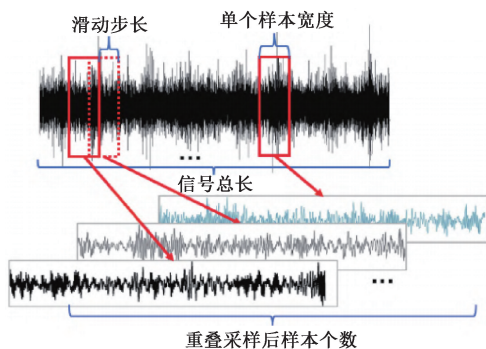


图 1 信号重叠采样

Fig. 1 Overlapping signal sampling

单个样本宽度 X_w 应不小于齿轮旋转一周的点数,即:

$$X_w \geq \frac{60}{n} \cdot F_s \quad (1)$$

式中: n 为转速(单位为 r/min); F_s 为采样频率(单位为 Hz)。

样本个数 N 可由式(2)算出,即:

$$N \leq \frac{L_t - X_w}{L_s} + 1 \quad (2)$$

式中:信号总长度为 L_t , 单个样本的宽度为 X_w , 每次滑动步长 L_s , 若重叠采样则 $L_s \leq X_w$ 。

1.2 短时傅里叶变换

短时傅里叶变换(short-time Fourier transform, STFT)是一种时频分析方法,能将非平稳的齿轮箱振动信号转换成反映信号时间变化和频率分布的二维时频图,可揭示故障引发的特定频率成分随时间的演变规律。

首先选择合适的窗函数 $w(t)$ (如汉宁窗、矩形窗等),窗长 L_w 需根据信号特性确定,窗长过短会导致频率分辨率下降,窗长过长则会降低时间分辨率,通常需兼顾两者以平衡时频分析效果,即:

$$L_w = \alpha \cdot \frac{F_s}{f_{\min}} \quad (3)$$

式中: α 为调节系数(一般取 2~5), $f_{\min}$ 为信号中需识别的最低特征频率。

然后对原始信号进行分帧加窗,拆成多个短时段:

$$x_m(t) = x(t) \cdot w(t - m \cdot L_s) \quad (4)$$

式中: m 为帧索引, L_s 为帧移。

将取到的短时片段进行傅里叶变换,得到该片段的频谱,STFT 的定义为:

$$STFT(m, k) = \sum_{t=0}^{L_w-1} x_m(t) \cdot e^{-j \cdot 2\pi \cdot k \cdot t / L_w} \quad (5)$$

式中: k 为频率索引。

将所有帧的频谱按时间顺序排列,构建二维矩阵。矩阵的行对应时间帧,列对应频率点,矩阵元素的幅值表示该时频点的能量强度,即形成 STFT 时频如图 2 所示。

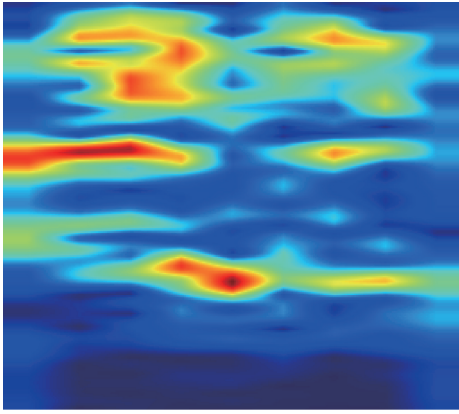


图 2 STFT 时频

Fig. 2 STFT time-frequency representation

2 故障诊断模型搭建

2.1 感受野注意力残差网络

感受野注意力残差网络是在经典残差网络基础上引入感受野注意力机制。该结构显著增强了对时频图中关键特征的提取能力。感受野注意力机制^[14]的核心是突破传统卷积的固定核约束,为感受野内各参数分配独立注意力权重,实现自适应特征提取。与空间注意力机制^[15]生成二维权重图不同,RFA 生成四维注意力张量,不仅能关注空间位置,还能区分感受野内部各点的重要性,实现更精细的“参数级”动态感受野调整。

本研究网络主体由多个残差块串联构成。每个残差块的结构如图 3 所示,其内部包含两个 3×3 卷积层,每层后接批归一化和 ReLU 激活函数,以有效提取局部特征并维持训练稳定性。为匹配输入与残差路径的通道数,使用 1×1 卷积进行维度调整。

为了增强 ResNet 的特征提取能力,本研究网络在多个层级嵌入感受野注意力机制,实现基于全局感受野的特征重标定,从而增强重要通道的响应,抑制噪声通道,显著提升模型的故障识别性能。该注意力机制如图 4 所示。

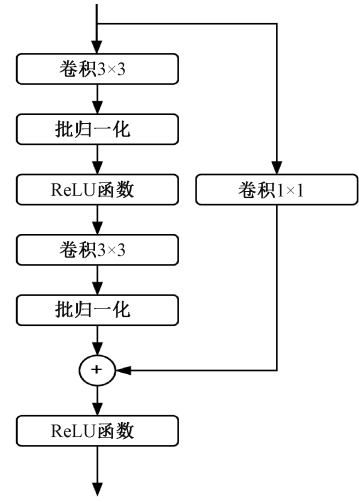


图 3 残差块结构

Fig. 3 Residual block structure

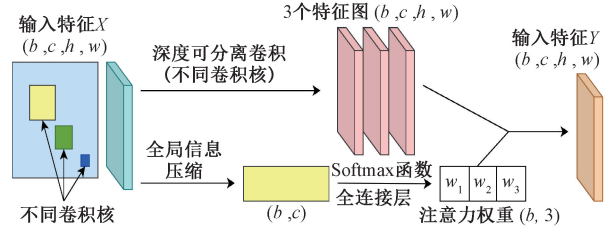


图 4 感受野注意力机制

Fig. 4 Receptive field attention mechanism

1) 四维注意力张量的生成:

首先,对输入特征图 $X \in \mathbf{R}^{b \times c \times h \times w}$ (其中 b 为批大小, c 为通道数, h, w 为高和宽)进行全局平均池化,压缩空间信息,得到通道级的全局上下文向量 $\mathbf{Z} \in \mathbf{R}^{b \times c}$ 。

$$\mathbf{Z} = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w X_c(i, j) \quad (6)$$

随后,将向量 $\mathbf{Z}$ 通过一个包含两个全连接层的瓶颈结构,并进行非线性激活。该变换的输出 $A_{raw} \in \mathbf{R}^{b \times c \times k \times k}$ 包含了重塑(Reshape)卷积核参数权重的信息。

$$A_{raw} = \mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \cdot \mathbf{Z} + b_1) + b_2 \quad (7)$$

式中: $\mathbf{W}_1, \mathbf{W}_2$ 为权重矩阵, b_1, b_2 为常数项。接着,将 A_{raw} 重塑为 $A_{k \times k} \in \mathbf{R}^{b \times c \times k \times k}$ 是四维注意力张量, k 为卷积核的高和宽。张量中的每一个元素 $A_{b,c,i,j}$ 代表第 b 个样本,第 c 个输出通道的卷积核在 (i, j) 位置上的参数的重要性权重。最后,通过 Sigmoid 函数将权重归一化。

$$A_{k \times k} = \text{Sigmoid}(\text{Reshape}(A_{raw})) \quad (8)$$

2) 多尺度特征提取与加权:用 3 个不同大小的深度可分离卷积(分组数等于输入通道数)对输入特征进行卷积,分别捕捉局部、中等和全局范围的上下文信息。在深度卷积阶段,将步骤 1)生成的四维注意力张量与对应的原始深度卷积核进行逐元素相乘,实现参数级的注意力加权,形成动态卷积核。

$$F_{3 \times 3} = X * (K_{3 \times 3} \odot A_{3 \times 3}) \quad (9)$$

$$F_{5 \times 5} = X * (K_{5 \times 5} \odot A_{5 \times 5}) \quad (10)$$

$$F_{7 \times 7} = X * (K_{7 \times 7} \odot A_{7 \times 7}) \quad (11)$$

式中: $K_{k \times k}$ 是原始卷积核权重, 形状与 $A_{k \times k}$ 相同, $\odot$ 表示逐元素相乘, $*$ 表示深度可分离卷积。

3) 加权融合多尺度特征: 将得到的多个尺度的特征图进行融合。融合权重 w_1, w_2, w_3 通过全连接层作用于全局上下文向量 $\mathbf{Z}$ 来实现, 并使用 Softmax 函数进行归一化, 确保权重之和为 1。

$$[w_1, w_2, w_3] = \text{Softmax}(\mathbf{W}_f \cdot \mathbf{Z} + b_f) \quad (12)$$

$$Y = w_1 F_{3 \times 3} + w_2 F_{5 \times 5} + w_3 F_{7 \times 7} \quad (13)$$

式中: $\mathbf{W}_f$ 为权重矩阵, b_f 为常数项。

2.2 基于优化的 BiGRU 网络

双向循环神经网络 (bidirectional recurrent neural network, BiRNN) 由学者在 1997 年提出。BiGRU 的本质就是将 GRU 单元用到 BiRNN 的结构框架中。GRU 通过引入更新门和重置门, 解决标准 RNN 难以捕捉长距离依赖和梯度消失/爆炸的问题。GRU 结构公式为:

$$z_t = \sigma(\mathbf{W}_z x_t + \mathbf{U}_z h_{t-1} + \mathbf{b}_z) \quad (14)$$

$$r_t = \sigma(\mathbf{W}_r x_t + \mathbf{U}_r h_{t-1} + \mathbf{b}_r) \quad (15)$$

$$\tilde{h}_t = \tanh(\mathbf{W}_h x_t + r_t \odot (\mathbf{U}_h h_{t-1}) + \mathbf{b}_h) \quad (16)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (17)$$

式中: σ 为 sigmoid 激活函数, $\odot$ 为逐元素乘法, $\mathbf{W}, \mathbf{U}$ 为权重矩阵, $\mathbf{b}$ 为偏置向量, h_{t-1} 为上一时刻的状态值, x_t 为当前时刻的输入值, r_t 和 z_t 分别为重置门和更新门, 通过 h_{t-1} 和 x_t 获取两个门控状态, $\tilde{h}_t$ 为当前时刻节点的信息, 首先通过通过重置门 r_t 对 h_{t-1} 进行重置, 再与 x_t 进行拼接, 最后通过 $\tanh$ 激活函数得到 $\tilde{h}_t$, h_t 为当前输出值, 通过更新门 z_t 实现遗忘和记忆。GRU 结构如图 5 所示。

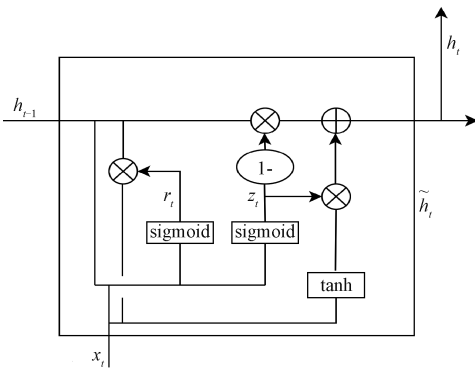


图 5 GRU 结构

Fig. 5 GRU architecture

GRU 只能根据过去的信息来推断当前状态, 而 BiGRU 则通过正向层和反向层的结合, 使得模型在理解序列中任何一个元素 (词、时间点、特征) 时, 不仅能参考它前面的元素 (历史信息), 还能参考它后面的元素 (未来信息),

相比单向 GRU 具有显著优势。BiGRU 网络模型如图 6 所示。

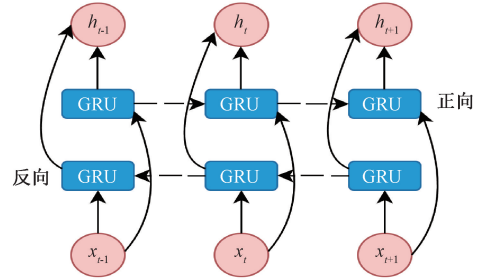


图 6 BiGRU 网络模型

Fig. 6 BiGRU network model

在 2017 年, 自注意力机制被 Vaswani 等^[16] 通过 Transformer 模型将其确立。自注意力机制让序列中的每个元素都与所有元素 (包括自身) 进行交互, 计算它们的相关性, 从而学习序列内部的依赖结构。BiGRU 通过双向结构捕获了序列的全局信息, 但无法区分关键信息与非关键信息, 而自注意力机制可以动态聚焦关键区域, 极大优化了 BiGRU 网络。注意力模块结构如图 7 所示。

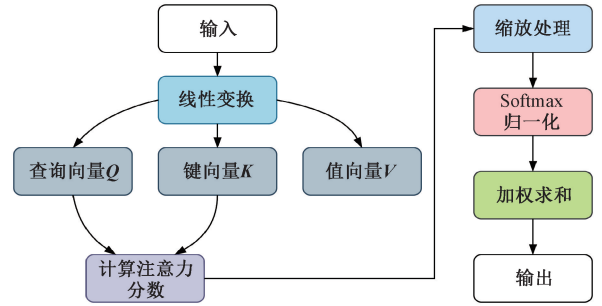


图 7 注意力模块结构

Fig. 7 Structure of the attention module

首先自注意力机制的核心公式为:

$$\mathbf{Q} = \mathbf{XW}_Q \quad (18)$$

$$\mathbf{K} = \mathbf{XW}_K \quad (19)$$

$$\mathbf{V} = \mathbf{XW}_V \quad (20)$$

$$\text{Attention} = \text{Softmax}\left(\frac{\mathbf{QK}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (21)$$

式中: $\mathbf{Q}$ 为查询矩阵; $\mathbf{K}$ 为键矩阵; $\mathbf{V}$ 为值矩阵, 用于进行加权平均; $\mathbf{QK}^T$ 为注意力分数, 实际上是每个 Query 和每个 Key 的相似度; d_k 为键矩阵的列数, 用于缩放处理以确保 Softmax 的输入在合理范围; Softmax 为归一化处理, 确保所有的权重都是正数且每行权重和为 1。

本文中, 自注意力机制被嵌入在 BiGRU 层之后, 对 BiGRU 输出的所有时间步隐藏状态进行加权融合, 从而突出关键时间节点对分类的贡献。具体结构如图 8 所示。

2.3 ResNet-RFAConv 与 BiGRU-SATT 模型结构

该模型是一个结合图像与信号处理优势的混合架构。

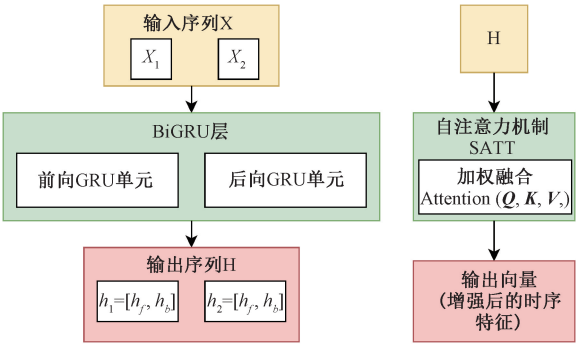


图 8 BiGRU-SATT 模块结构

Fig. 8 BiGRU-SATT modular structure

模型通过异构分支分别处理图像与信号,充分发挥不同数据的互补性。图像分支基于改进的 ResNet 结构,通过 RFA 注意力机制和残差块分层提取视觉特征,最终生成 512 维特征;信号分支融合 BiGRU 与自注意力机制, BiGRU 捕捉时序双向依赖,自注意力机制聚焦重要特征,协同输出 256 维特征。图像分支引入 RFA 模块增强局部特征,信号分支结合自注意力机制提升时序建模能力。两分支特征拼接后,经全连接层降维、ReLU 激活和 Dropout 正则化,最终实现分类输出。ResNet-RFA + BiGRU-SATT 网络如图 9 所示。

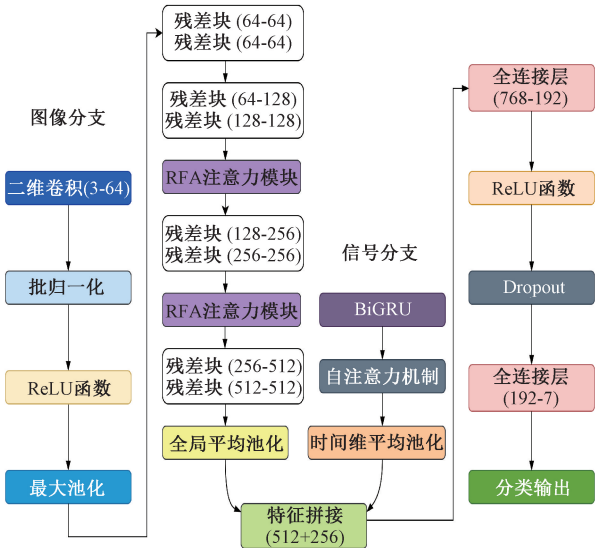


图 9 ResNet-RFA+BiGRU-SATT 网络结构

Fig. 9 ResNet-RFA+BiGRU-SATT network architecture

2.4 弧齿锥齿轮故障诊断模型

基于 ResNet-RFA + BiGRU-SATT 的齿轮箱智能故障诊断流程如图 10 所示,诊断步骤包括:

- 1)对原始信号进行重叠采样以增加样本个数。
- 2)将样本数据集划分成训练集和测试集。
- 3)将所有的样本进行短时傅里叶变换,将样本信号由一维振动数据转化为 256×256 的二维 RGB 时频图。

- 4)将所有样本增加维度使振动信号转化为时序信号。
- 5)将图和信号进行训练,并保存准确率最大的参数。
- 6)使用训练好的 ResNet-RFA + BiGRU-SATT 模型对弧齿锥齿轮箱故障数据集进行分类,输出诊断结果。

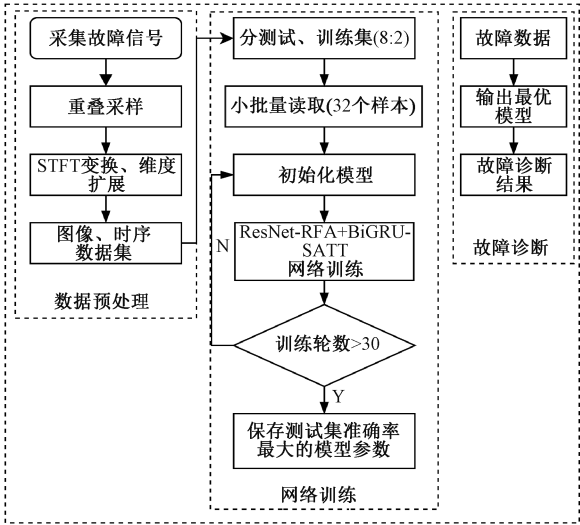


图 10 弧齿锥齿轮故障诊断模型

Fig. 10 Fault diagnosis model for spiral bevel gears

3 弧齿锥齿轮箱故障模拟实验

3.1 模拟故障实验及数据集

故障模拟实验在晋中学院故障诊断实验室进行,实验台由驱动电机、扭矩传感器、ICP 单向加速度传感器、ICP 三向加速度传感器、弧齿锥齿轮箱、行星齿轮减速器、底板、磁粉制动器、防护罩、控制系统和采集仪等部分组成。安装位置关系如图 11 所示。

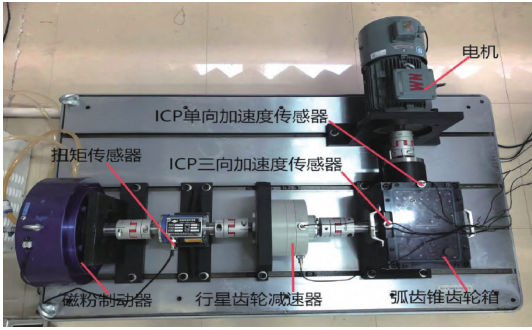


图 11 弧齿锥齿轮故障诊断实验台

Fig. 11 Spiral bevel gear fault diagnosis test rig

加速度传感器 X、Y、Z 轴分别对应采集通道 1、2、3, ICP 单向加速度传感器 Z 轴对应信号采集通道 4,传感器都布置在齿轮轴线的箱盖上,测点位置如图 12 所示。

本实验电机转速 1 000 r/min,采样频率 5 000 Hz,采用 DASP 数据采集系统采集不同故障工况的振动加速度信号,采集 7 种工况,分别为正常工况、大齿轮缺齿、小齿轮缺

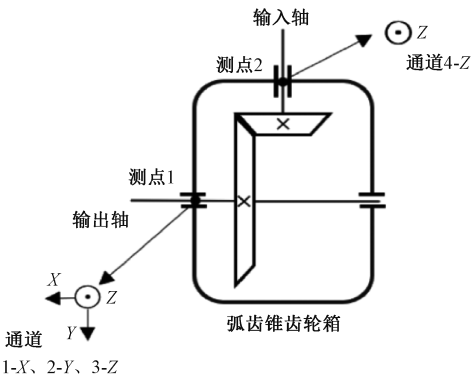


图 12 测点位置

Fig. 12 Measurement location

齿、大齿轮齿根裂纹、小齿轮齿根裂纹、大齿轮齿面磨损、小齿轮齿面磨损,每种工况采样时间 60 s。由式(1)取单个样本长度 1 024,设置步长为 128,每种工况样本数为 1 120,7 种工况共 7 840 个样本。设置训练集和测试集的比例为 8:2,训练样本数量为 6 272,测试样本数量为 1 568。本研究在提取特征时,采用 STFT,通过引入滑动时间窗将非平稳的一维时间序列信号分割为多个近似平稳的短时信号片段,再对每个片段进行傅里叶变换以获取局部频谱特征,最终将各片段的频谱按时间顺序排列,形成 256×256 的时频图。各工况时频图如图 13 所示。

3.2 实验结果与分析

为了全面检验模型性能,将实验室数据集输入到 ResNet-

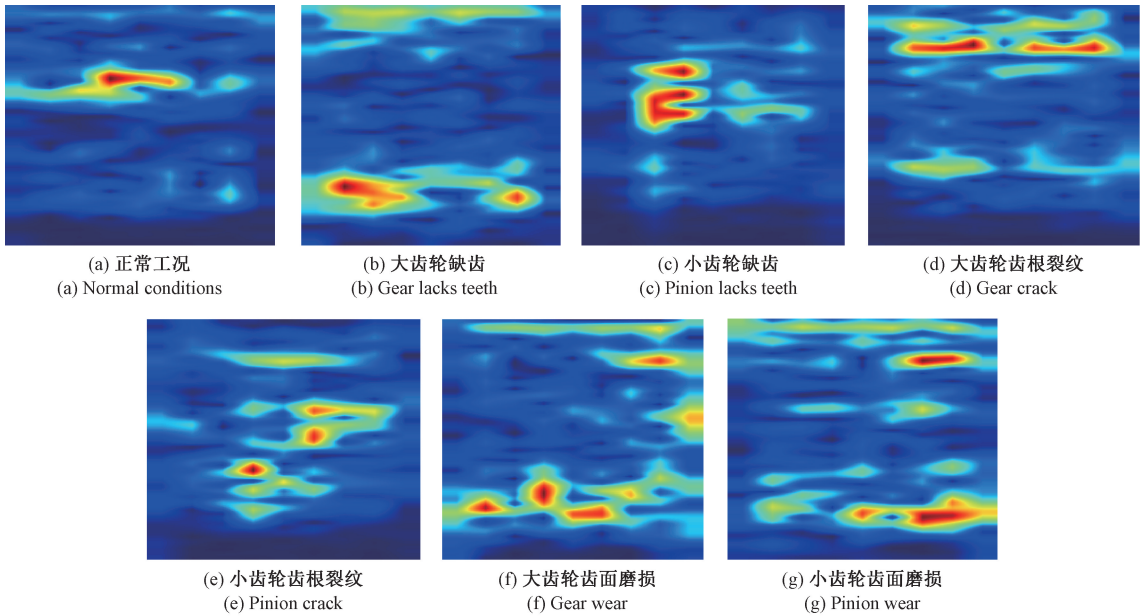


图 13 时频图示例

Fig. 13 Example of a time-frequency representation

RFA+BiGRU-SATT 混合架构中进行多次训练。初始学习率为 0.001,动态调整学习率以获得最优参数,dropout

值选取 0.5,以防止过拟合现象,模型经过 30 轮训练结果如图 14 所示。

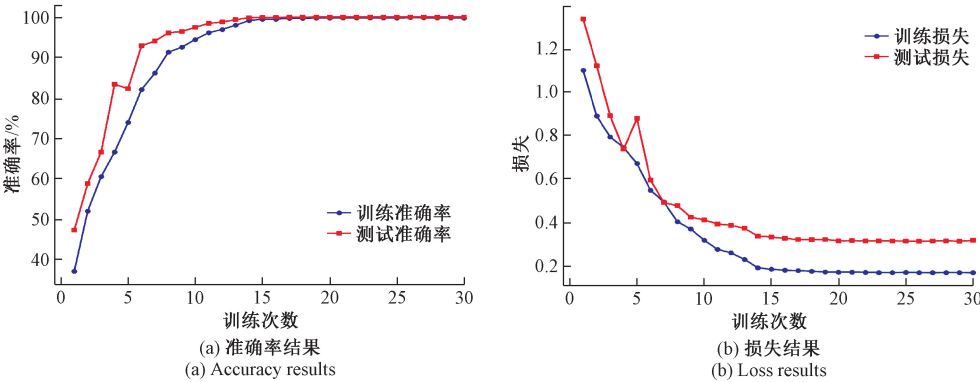
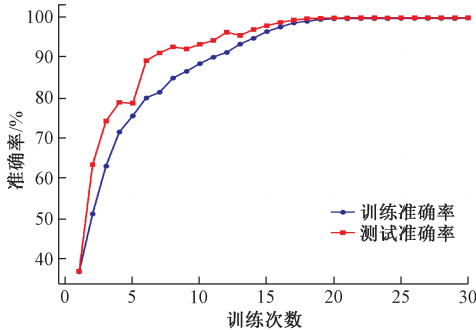


图 14 模型准确率及损失值变化

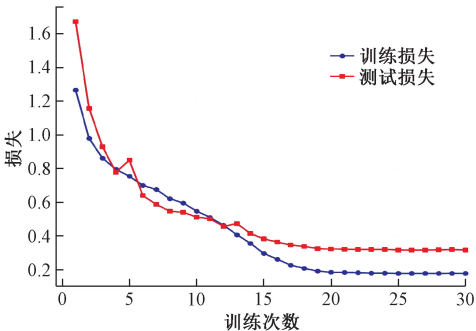
Fig. 14 Variations in model accuracy and loss curves

模型训练与测试的准确率随训练次数增加均呈稳步上升趋势,训练损失与测试损失随训练次数增加均呈持续下降趋势,最终趋于平稳;且测试损失与训练损失的差距极小,表明模型在优化过程中,不仅能有效降低训练集的预测误差,对未参与训练的测试集,误差也能同步降低并稳定在低水平,验证了模型的泛化稳定性。

为了进一步评估模型在真实工业场景下的鲁棒性与实用性,向原始振动信号中添加了信噪比为 10、20 dB 的高斯-脉冲混合噪声(脉冲出现概率为 5%)以模拟实际工况,高斯噪声模拟了环境中的随机干扰,而脉冲噪声则对应设备运行中可能出现的突发性扰动。添加 10、20 dB 噪声的模型准确率和损失变化如图 15、16 所示。



(a) 准确率结果
(a) Accuracy results



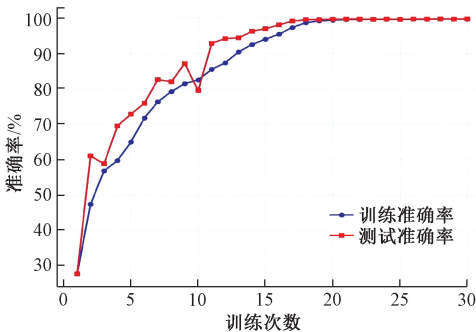
(b) 损失结果
(b) Loss results

图 15 添加 10 dB 噪声的模型准确率及损失值变化

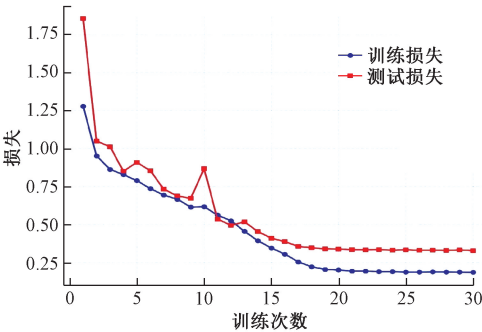
Fig. 15 Changes in model accuracy and loss with added 10 dB noise

对于添加 10 dB 噪声的模型,从准确率曲线观察,训练准确率与测试准确率同步上升,在约 20 次训练迭代后几乎重合且趋近于 100%,这表明模型在 10 dB 噪声干扰下,不仅学习能力出色,且能很好地将学到的知识迁移到未见过的带噪数据上,泛化性能优异。从损失曲线来看,训练损失与测试损失均呈快速下降趋势,在 20 次训练后稳定在 0.2 左右的低区间,二者全程差距微小,说明模型在拟合与泛化的一致性上表现突出。

对于添加 20 dB 噪声的模型,准确率曲线中测试准确率在训练过程中出现明显波动,例如在第 10 次训练附近有显著下降,且训练准确率与测试准确率的收敛速度也慢



(a) 准确率结果
(a) Accuracy results



(b) 损失结果
(b) Loss results

图 16 添加 20 dB 噪声的模型准确率及损失值变化

Fig. 16 Changes in model accuracy and loss with added 20 dB noise

于 10 dB 噪声场景。从损失曲线分析,测试损失的波动更为剧烈,同样在第 10 次训练左右出现大幅上升,整体损失下降的效率更低,最终稳定值也高于 10 dB 噪声下的结果,反映出 20 dB 噪声给模型学习带来了更多干扰,增加了其拟合难度。

整体而言,噪声分贝的高低对模型性能有着显著影响,分贝越高,模型学习的“阻碍”就越强。具体表现为准确率的收敛速度变慢、训练与测试的损失下降过程更曲折、收敛速度更缓且最终稳定值更高。不过,即便在 20 dB 的高噪声环境下,模型的训练准确率和测试准确率最终仍能达到较高水平,损失也能稳定在较低范围,这说明该模型具备一定的抗噪声鲁棒性,能够在噪声干扰下学习到数据的有效特征。在实际应用中若面临强噪声场景,可能需要引入降噪预处理步骤来提升模型性能。

为了进一步展示实验结果,本研究绘制了无噪声、加入 10、20 dB 噪声时模型的混淆矩阵如图 17 所示。

对于无噪声的混淆矩阵(图 17(a)),所有真实标签与预测标签的交叉处均只有对角线上有数值(且为 1 120),无其他误分类情况,说明模型在无噪声环境下分类完全准确,泛化能力极强。对于加入 10 dB 噪声的混淆矩阵(图 17(b)),大部分类别仍能准确分类,出现了少量误分类,不过整体仍保持较高的分类准确性。对于加入 20 dB 噪声的混淆矩阵(图 17(c)),误分类情况进一步增加,误分

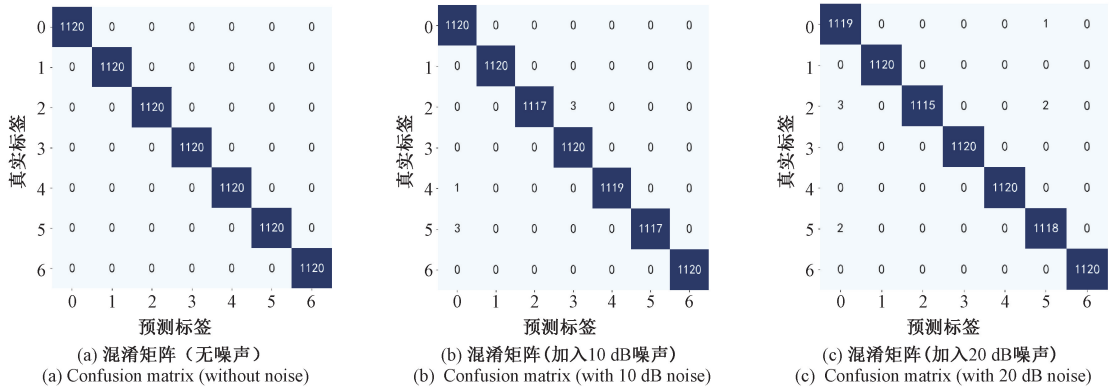


图 17 不同分贝噪声下模型的混淆矩阵对比

Fig. 17 Comparison of model confusion matrices under different decibel noise levels

类的数量和涉及类别都比 10 dB 噪声时更多。

整体来看,噪声会对模型的分类性能产生负面影响,且噪声分贝越高,对模型分类准确性的干扰越大,但即使在 20 dB 噪声下,模型对大部分类别仍能保持准确分类,说明该模型在一定程度上具备抗噪声的鲁棒性。

3.3 计算复杂度与实时性分析

为评估模型在实际工业场景中的可行性,本研究从计算复杂度和推理效率两方面对 ResNet-RFA + BiGRU-SATT 模型进行了分析。理论分析表明,模型计算负荷主要来自双分支结构:ResNet-RFA 分支以卷积和感受野注意力机制为主,其复杂度与时频图尺寸和通道数相关;BiGRU-SATT 分支则依赖于时序长度和隐层维度,其中自注意力机制随序列长度增长而计算量增加,但在本任务设定的序列长度下仍属可接受范围。在 NVIDIA RTX 3080 GPU 平台上的实测结果表明,模型单次推理时间约为 12.3 ms,参数量为 18.7 M,运行内存占用约 320 MB,表现出较高的效率。

本研究进一步探讨了模型在工业环境中的适用性。尽管模型在噪声环境下仍保持高精度,但其参数量和推理时间在资源受限的边缘设备上可能成为部署瓶颈。为此,建议在实际应用中采取多种轻量化策略,如对输入时频图进行降采样,或引入模型剪枝与量化技术,以进一步提升推理速度并降低资源消耗。综上所述,该模型在中等算力平台上已具备良好的部署潜力,未来将通过结构优化与硬件适配进一步提升其实时性能与工业适用性。

3.4 模型稳定性与统计显著性分析

为全面评估模型性能并排除过拟合风险,本研究进行了深入的统计显著性分析:

1)多次重复实验:在相同实验条件下重复运行模型 30 次,每次随机划分训练集和测试集。结果显示,模型平均准确率为 $99.93\% \pm 0.12\%$,表明模型具有优异的稳定性。模型分类准确率的分布箱型图如图 18 所示。

2)平均混淆矩阵:基于 30 次实验计算平均混淆矩阵,对角线元素平均值达 99.9% 以上,标准差 $< 0.12\%$,各类

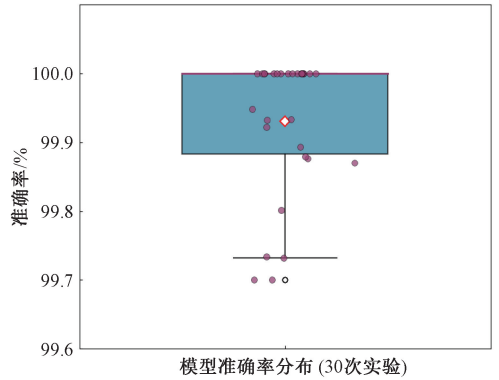


图 18 模型分类准确率的分布

Fig. 18 Distribution of classification accuracy

别分类性能均稳定可靠。具体而言,每个类别的准确率均保持在 99.85%~99.92% 之间,而误分类率极低(所有非对角线元素均低于 0.04%),进一步验证了模型在各类别上的分类可靠性。平均混淆矩阵如图 19 所示。

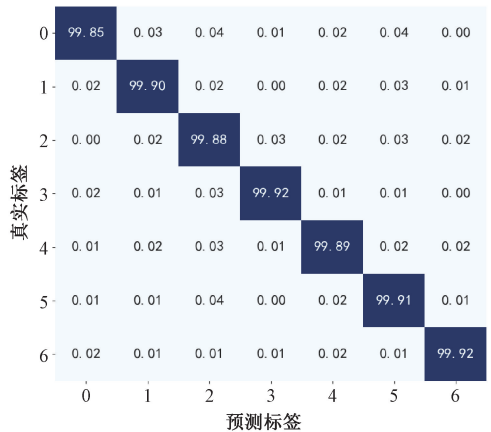


图 19 平均混淆矩阵

Fig. 19 Average confusion matrix

3)交叉验证分析:采用 5 折交叉验证方法对模型进行评估。结果显示,各折准确率均高于 99.8%,平均准确率

达 99.92%,标准差仅为 0.000 4。这表明模型具有优异的稳定性和泛化能力,对数据划分不敏感,有效避免了过拟合现象。 K 折交叉验证准确率分布如图 20 所示。

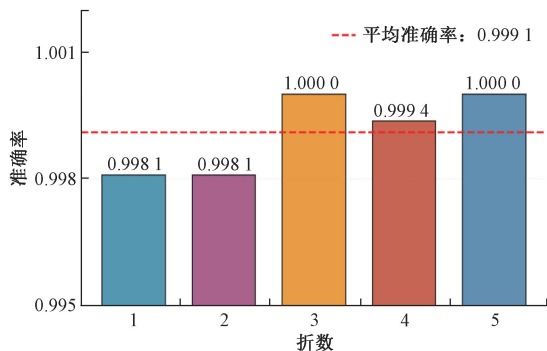


图 20 K 折交叉验证准确率分布

Fig. 20 K -fold cross-validation accuracy distribution

3.5 不同网络模型的结果对比

为进一步探究所提 ResNet-RFA + BiGRU-SATT 模型的效果,以及不同网络结构对任务性能的影响,分别设计 Transformer、Mamba 新兴模型和 SVM、DBN 经典模型作为对比,探究各类网络对当前任务的适应性,从而更全面地验证本文模型的优势与合理性。如图 21 所示。

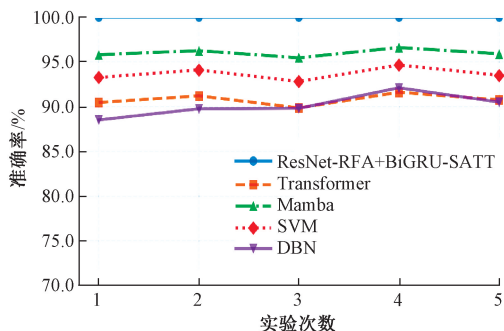


图 21 不同网络模型对比结果

Fig. 21 Comparison of different network models

对比结果显示:ResNet-RFA + BiGRU-SATT 模型取得了最优性能,准确率接近 100%,且表现出最佳的稳定性。在对比模型中,Mamba 架构表现最佳(96%),传统 SVM 方法(94%)也展现出良好性能,甚至优于 Transformer(91%),而 DBN(89%)则相对较弱。总体来看,本文模型在准确率与稳定性方面均显著优于所有对比模型;同时,多数深度学习模型展现出强大竞争力,而传统方法中 SVM 虽表现尚可,但整体上仍落后于先进的深度学习架构。

4 结 论

本研究提出一种 ResNet-RFA + BiGRU-SATT 的融合模型,并通过对比实验验证其性能,得出以下结论:

该模型在弧齿锥齿轮箱故障诊断任务中表现出卓越

性能,其准确率高、稳定性强,经多次实验验证始终保持接近 100% 的识别准确率,充分证明了该融合架构的有效性与可靠性。具体而言,ResNet-RFA 模块通过残差结构实现了深层特征的有效提取与复用,增强了空间特征的判别能力;BiGRU-SATT 模块则借助双向门控机制捕捉序列数据的时序依赖关系,并结合自注意力机制进一步聚焦关键时间节点,提升了对动态特征的建模精度。二者的有机结合,实现了“特征提取—时序建模—关键聚焦”的高效协同,使模型在复杂噪声环境下仍保持优异的鲁棒性。

与其他模型相比,本研究所提融合模型在准确率、稳定性及模块协同效率等方面均表现更优,不仅在无噪声条件下达到完全识别,在高斯-脉冲混合噪声干扰下仍具备较高的分类性能,展现出较强的工程适用性。

参考文献

- [1] 苏缪诞. 经验模态分解改进算法及其在旋转机械故障诊断中的应用研究[D]. 马鞍山:安徽工业大学, 2020. SU M X. Improved algorithm of empirical mode decomposition and its application in rotating machinery fault diagnosis[D]. Ma'anshan: Anhui University of Technology, 2020.
- [2] 张建峰, 丁传仓, 江星星, 等. 基于加权双域稀疏分解模型的编码器信号特征提取及其行星齿轮箱故障诊断应用[J]. 仪器仪表学报, 2025, 46(7): 260-270. ZHANG J F, DING CH C, JIANG X X, et al. Encoder signal feature extraction based on weighted dual-domain sparse decomposition model and its application in planetary gearbox fault diagnosis[J]. Chinese Journal of Scientific Instrument, 2025, 46(7): 260-270.
- [3] 别锋锋, 张瀚阳, 李倩倩, 等. 基于多元变分模态分解的行星齿轮箱故障诊断方法[J]. 机械设计与研究, 2025, 41(2): 364-370. BIE F F, ZHANG H Y, LI Q Q, et al. Fault diagnosis method for planetary gearboxes based on multivariate variational mode decomposition[J]. Machine Design & Research, 2025, 41(2): 364-370.
- [4] GUO SH, YANG T, GAO W, et al. A novel fault diagnosis method for rotating machinery based on a convolutional neural network[J]. Sensors, 2018, 18(5): 1429.
- [5] 赵剑平. 基于深度置信网络的齿轮箱故障诊断方法研究[D]. 北京:中国石油大学, 2019. ZHAO J P. Fault diagnosis method of gearbox based on deep belief network[J]. Beijing: Journal of China University of Petroleum, 2019.
- [6] WANG L Y, ZHAO W G. An ensemble deep learning network based on 2D convolutional neural network and

- 1D LSTM with self-attention for bearing fault diagnosis [J]. Applied Soft Computing, 2025, 172: 112889.
- [7] 张雄,渠伟滢,王文强,等. 基于改进 CNN-GRU 模型的滚动轴承多故障诊断模型[J]. 机电工程, 2025, 42(10):1931-1939.
- ZHANG X, QU W Y, WANG W Q, et al. Multi-fault diagnosis model for rolling bearings based on improved CNN-GRU model[J]. Journal of Mechanical & Electrical Engineering, 2025, 42(10): 1931-1939.
- [8] 荀小伟,许昕,潘宏侠. 基于 DCNN 和 Bi-LSTM 的弧齿锥齿轮箱故障诊断[J]. 电子测量技术, 2024, 47(10): 48-55.
- XUN X W, XU X, PAN H X. Fault diagnosis of spiral bevel gearbox based on DCNN and Bi-LSTM [J]. Electronic Measurement Technology, 2024, 47(10): 48-55.
- [9] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [10] 魏秀业,程海吉,贺妍,等. 基于特征融合与 ResNet 的行星齿轮箱故障诊断[J]. 电子测量与仪器学报, 2022, 36(5): 213-222.
- WEI X Y, CHENG H J, HE Y, et al. Fault diagnosis of planetary gearbox based on feature fusion and ResNet[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(5): 213-222.
- [11] 韩延,李超,黄庆卿,等. 小样本下时序注意力边界增强原型网络的齿轮箱故障诊断方法[J]. 电子测量与仪器学报, 2023, 37(2): 90-98.
- HAN Y, LI CH, HUANG Q Q, et al. Fault diagnosis method of gearbox based on temporal attention boundary enhanced prototype network under small sample condition [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37 (2): 90-98.
- [12] 王翰伟,许昕,潘宏侠,等. 基于 GADF 和 HWP-CBAM-ResNet 的弧齿锥齿轮箱故障诊断[J]. 电子测量技术, 2024, 47(21): 100-110.
- WANG H W, XU X, PAN H X, et al. Fault diagnosis of spiral bevel gearbox based on GADF and HWP-CBAM-ResNet [J]. Electronic Measurement Technology, 2024, 47(21): 100-110.
- [13] 吴启航,丁晓喜,何清波,等. 齿轮箱故障边缘智能诊断方法及应用研究[J]. 仪器仪表学报, 2024, 45(1): 70-80.
- WU Q H, DING X X, HE Q B, et al. Edge intelligent diagnosis method and application for gearbox faults [J]. Chinese Journal of Scientific Instrument, 2024, 45(1): 70-80.
- [14] LIU S T, HUANG D, WANG Y H. Receptive field block net for accurate and fast object detection[C]. European Conference on Computer Vision, 2018: 404-419.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. ArXiv preprint arXiv: 1706.03762,2017.

作者简介

王晓晨, 硕士研究生, 主要研究方向为机械制造、过程装备运行状态监测与故障诊断。

E-mail:18835418388@163.com

许昕(通信作者), 副教授, 主要研究方向为故障诊断、信号处理。

E-mail:ninaxx79@163.com