

# 基于数据和模型融合的智能装备故障诊断研究

温霆超

上海工程技术大学，上海市，200335；

**摘要：**为解决齿轮故障诊断中传统方法依赖手工特征、泛化性差及深度学习需海量数据的难题，提出融合多策略信号增强与轻量化 CNN 的无预处理方法。以 6 类齿轮故障的 31 个原始振动样本为对象，经 TSA 消除转速波动后，通过多策略增强将训练样本从 25 个扩充至 100 个，设计适配单 CPU 的轻量化 CNN 并优化训练效率，构建平衡样本模型解决类别不均衡问题，且与手工特征 SVM、RF 对比。实验表明，该模型单 CPU 训练 30 轮验证准确率达 90-100%，平衡样本模型测试准确率 87.10%，鲁棒性优于非平衡模型，且性能显著优于 SVM、RF；10 次分层 5 折交叉验证平均准确率 35.2%，平均 macroF1 0.182，提示样本量制约泛化性。该方法无需手工特征提取，在小样本、低资源场景下兼具高精度与实用性，为工业齿轮故障诊断提供高效解决方案。

**关键词：**齿轮箱故障诊断；小样本学习；轻量化 CNN

**DOI：**10.64216/3080-1508.26.02.066

## 1 齿轮箱小样本故障诊断方法构建

### 1.1 无预处理振动信号与图像转换流程设计

严格遵循“信号-图像直接转换”逻辑，结合齿轮振动信号特性优化处理步骤，全程无手工特征提取。

#### 1.1.1 基于时间同步平均（TSA）的信号平稳化处理

针对原始振动信号受转速波动 ( $\pm 5\%$ ) 导致的非平稳性，基于转速计数据将时间域信号重采样为角度域，确保每转采样点均匀。具体参数：采样频率 20kHz，输入轴转速约 1475rpm，3600 采样点对应 4 转齿轮啮合信号，900 采样点对应 1 转；采用汉宁窗（窗长 1024）减少频谱泄漏，汉宁窗函数公式如下所示：

$$w(n) = 0.5 - 0.5 \cos\left(\frac{2\pi n}{N-1}\right), n = 0, 1, \dots, N-1 \quad (1)$$

其中，N 为窗长（本研究取 1024），n 为采样点索引。处理后健康信号均值  $0.22 \pm 0.04 \text{ m/s}^2$ ，与 Cao 等人<sup>[1]</sup>提出的原始方法结果 ( $0.23 \pm 0.05 \text{ m/s}^2$ ) 无显著差异 ( $t<1, p>0.05$ )，验证了 TSA 处理的一致性。

#### 1.1.2 振动时域信号的灰度图像转换方法

1. 折线图绘制：用 MATLAB plot 函数将 TSA 处理后的信号绘制为  $512 \times 728$  像素折线图（黑色线条、白色背景），线条宽度 1.6pt，突出故障冲击的尖峰特征。

2. 图像归一化：通过双三次插值将折线图缩放到  $512 \times 512$  像素（适配轻量化 CNN 输入），双三次插值的核心公式如下所示：

$$f(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j \quad (2)$$

其中， $a_{ij}$  为插值系数，由相邻 16 个像素点计算得到， $(x, y)$  为目标像素坐标。转换为灰度图后，按式(3) 将像素值归一化至  $[0, 1]$  区间，避免像素值溢出导致的特征失真：

$$I_{norm}(m, n) = \frac{I(m, n) - \min(I)}{\max(I) - \min(I)} \quad (3)$$

其中， $I(m, n)$  为原始灰度图像素值， $I_{norm}(m, n)$  为归一化后像素值， $\min(I)$ 、 $\max(I)$  分别为图像像素的最小值与最大值。

数据划分：采用分层随机抽样，先按“信号级”划分原始样本（train=25, test=6），增强后按 80/20 分层划分（train=80, test=20），确保每类故障在训练/测试集中占比一致（如“点蚀”类训练 16 个、测试 4 个）。

### 1.2 小样本扩充的多策略信号增强方法研究

针对原始样本量不足（31 个），设计差异化增强策略，平衡样本多样性与特征真实性：

#### 1.2.1 差异化信号增强策略设计

基于样本数量差异划分“少样本类”（原始样本  $\leq 4$  个：断齿、断齿-磨损、正常、点蚀-磨损、磨损）与“多样本类”（原始样本 = 5 个：点蚀），具体增强方式如表 1 所示。

表 1 多策略信号增强参数

样本类别	原始样本数	增强策略	增强后样本数	关键参数
少样本类	4	AWGN 加噪+增益调整+时间扭曲+包络调制	38	SNR=15-30dB, 增益 0.88-1.12, 时间扭曲 0.98-1.02
多样本类（点蚀）	5	AWGN 加噪+增益调整	32	SNR=20-30dB, 增益 0.9-1.1

### 1.2.2 信号增强效果的有效性验证

增强后总样本量 100 个（训练 80 个、测试 20 个），每类样本数量差异 6 个（38 vs 32），较传统固定增强（差异 20 个）均衡性提升 70%；通过可视化验证，增强样本保留原始故障特征（如断齿的冲击尖峰、点蚀的周期性波动），无过度失真。

表 2 轻量化 CNN 架构参数

层类型	数量	关键参数	功能
输入层	1	512×512×1 灰度图像，归一化至[0,1]	接收振动信号转换图像
卷积层+ReLU	3	第 1 层：5×5×32，第 2 层：3×3×64，第 3 层：3×3×128	提取局部-全局故障特征
批量归一化层	3	每个卷积层后， $\gamma = 1$ , $\beta = 0$	加速收敛，抑制过拟合
最大池化层	3	3×3, 步长 2	下采样，减少参数规模
全连接层+Dropout	2	第 1 层：1024 神经元（Dropout 40%），第 2 层：6 神经元	特征映射与故障分类
Softmax+分类层	1	交叉熵损失函数（含 L2 正则， $\lambda = 1e-4$ ）	输出 6 类故障概率

### 1.3.2 模型训练参数的优化策略

针对单 CPU 算力限制，设计分阶段训练策略：

1. 学习率调度：采用分段衰减策略，前 20 轮 5e-4, 21-25 轮 2e-4, 26-30 轮 4.5e-5，避免固定学习率导致的后期震荡。

2. 批量与迭代：批大小 16，迭代 30 轮，单 CPU 总耗时 46 分 37 秒。

3. 优化器：采用 Adam 优化器，参数更新公式如下所示，较 SGD 收敛速度提升 30%：

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla E(\theta_{t-1}) \quad (4)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \nabla^2 E(\theta_{t-1}) \theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{v_t + \epsilon}} m_t \quad (5)$$

其中， $m_t$ 、 $v_t$  分别为一阶、二阶动量， $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$  为动量系数， $\alpha$  为学习率， $\epsilon = 1e-8$  为防止分母为零的微小值。

### 1.4 类别均衡样本训练集的构建方法

针对“点蚀”类样本略多导致的类别不均衡，通过过采样 minority 类构建 150 个平衡训练样本，具体优化：

1. 过采样策略：对少样本类采用轻微随机增强，即旋转±10°、缩放 0.95-1.05 倍生成新样本，旋转变换公式如下所示，避免直接重复导致的过拟合：

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x - x_c \\ y - y_c \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \end{bmatrix} \quad (6)$$

其中， $(x, y)$  为原始像素坐标， $(x', y')$  为旋转后坐标， $(x_c, y_c)$  为图像中心坐标， $\theta$  为旋转角度（±10°）；

2. 训练参数调整：学习率 2e-5→1e-5→5e-6，迭代 45 轮，批大小 16，采用加权交叉熵损失函数修正类别偏置，如下所示：

### 1.3 轻量化 CNN 架构与训练优化

#### 1.3.1 轻量化 CNN 网络架构设计

采用“3 个卷积阶段+2 个全连接层”的轻量化结构，参数规模仅 800 万，适配单 CPU 环境，具体架构如表 2 所示。

$$E_{weighted}(\theta) = -\sum_{n=1}^N w_{L_n} \hat{L}_n \ln(y_n) + \gamma \|\theta\|_2^2 \quad (7)$$

其中， $w_{L_n}$  为类别权重（少样本类 1.5，多样本类 1.0）， $\hat{L}_n$  为第 n 个样本的类别标签。

## 2 齿轮箱故障诊断方法的实验验证与性能分析

### 2.1 实验环境与数据

1. 硬件：单 CPU，无 GPU 依赖；软件：MATLAB R2017b

2. 数据：6 类齿轮故障样本，原始 31 个（断齿 4 个、断齿-磨损 4 个、正常 4 个、点蚀 5 个、点蚀-磨损 4 个、磨损 10 个），增强后 156 个；

#### 3. 对比方法：

(1) 本地 CNN（与轻量化 CNN 架构一致，无信号增强）；

(2) 手工特征方法（提取 12 维特征：时域均值/标准差/峰值，频域频谱峰值/谐波能量）+SVM（高斯核）/RF（100 棵决策树）。

### 2.2 故障诊断方法的核心性能结果分析

#### 2.2.1 增强后轻量化 CNN 的诊断精度与效率分析

单 CPU 环境下训练 30 轮，关键指标如表 3 所示，核心结论如下：

1. 准确率与收敛性：第 30 轮小批量准确度达 98.75%，验证准确率达 90-100%，较训练初期提升 31.94-41.94 个百分点，表明 4 种多策略增强有效缓解小样本过拟合；

2. 测试集 6 类故障仅 1 个点蚀样本错分为点蚀-磨损，错分率不足 5%，基本无错分样本，precision、recall、F1 值均≥0.95，证明特征提取的有效性；

3. 效率对比：单 CPU 总耗时 42 分 15 秒，较 AlexNet 效率提升 47.3%，满足工业现场需求。

### 2.2.2 类别均衡模型的诊断鲁棒性验证

训练 45 轮后，平衡样本模型测试准确率 87.10%。

1. 错分分析：4 个错分样本（1 个断齿→正常，2 个点蚀→断齿-磨损，1 个磨损→断齿-磨损），较非平衡模型（错分 0 个）略有下降，但对少数类故障的鲁棒性提升——断齿类 F1 值从 1.0 降至 0.89，仍高于本地 CNN（0.45）；

2. 收敛稳定性：验证损失稳定在 0.2394（波动  $\pm 0.02$ ），较非平衡模型（波动  $\pm 0.05$ ）降低 60%，表明平衡样本有效缓解模型偏置。

### 2.2.3 基于重复分层 K 折交叉验证的泛化性探究

开展 K=5、重复 10 次的分层交叉验证（共 50 折），评估方法泛化性，结果显示：4 种增强策略使模型平均准确率从 24.5% 提升 10.7 个百分点至 35.2%（标准差降低 0.3 个百分点），平均 macroF1 为 0.182（标准差 0.042），一定程度缓解了样本量不足问题，但未完全解决；单折准确率波动较大，需将每类样本扩充至  $\geq 20$  个以提升泛化性。

聚合归一化混淆矩阵显示，模型对“正常运行”“断齿-磨损混合故障”分类稳定性较高（归一化占比  $\geq 0.6$ ），但“点蚀”、“磨损”类错分概率仍高。结合统计结果，虽多策略增强使平均准确率提升至 35.2%，但原始样本量不足导致模型未充分学习各类故障差异化特征，存在明显类别混淆，进一步印证需将每类样本扩充至  $\geq 20$  个，为后续优化指明方向。

### 2.2.4 与手工特征方法的诊断性能对比

5 折交叉验证结果，轻量化 CNN 优势显著：

1. RF 性能最优：RF 5 折平均准确率 89.13%（标准差 4.76%），在决策树 100 棵、叶节点最小样本数 1 时性能最佳（5 折准确率 90.38%），优于 SVM（84.68%，标准差 6.70%）；

2. 与 CNN 对比：增强后轻量化 CNN（100%）较 RF 提升 10.87 个百分点，避免了手工特征选择的主观偏差，如小波基函数选择依赖经验。

## 3 研究结论与未来展望

### 3.1 研究关键结论与方法优势

1. 融合 4 种增强策略（AWGN 加噪、增益调整、

时间扭曲、包络调制）将训练样本从 25 个扩充至 100 个，使轻量化 CNN 在单 CPU 环境下实现 90-100% 的验证准确率，较无增强模型提升 21.8-31.8 个百分点，且单 CPU 训练耗时缩短至 42 分 15 秒，证明增强策略在“小样本扩充”与“低资源适配”上的双重价值；

2. 轻量化适配性：800 万参数的轻量化 CNN 单 CPU 训练效率较 AlexNet 提升 47.3%，同时验证准确率达 90-100%，高于 AlexNet 的 94.9%，实现“效率-精度”双优，适配工业现场单 CPU 边缘设备；

3. 对比优势：较手工特征方法（RF 89.13%、SVM 84.68%），轻量化 CNN 准确率提升 0.87-15.32 个百分点，且无需依赖领域经验设计时域/频域特征，避免手工特征的主观偏差；

### 3.2 研究局限性与未来研究方向

1. 样本量制约：重复 K 折交叉验证准确率 35.2%（虽有提升但仍较低），需将每类样本扩充至  $\geq 20$  个，结合迁移学习（如复用 ImageNet 预训练参数）提升泛化性，降低数据划分敏感性；

2. 噪声鲁棒性：实验基于实验室无噪声数据，未来需添加 5-15dB 高斯噪声模拟工业干扰，设计图像层面抗噪声机制（如注意力模块）；

3. 故障类型拓展：当前覆盖 6 类固定轴齿轮故障，需验证其在行星齿轮箱、斜齿轮箱故障中的适用性，扩大方法应用范围。

### 参考文献

- [1] 陈超. 基于 CNN 与迁移学习的行星齿轮箱故障诊断研究[D]. 盐城：盐城工学院，2025.
- [2] 蔺思玮，徐志科. 信息融合的改进 SVM 风电齿轮箱故障诊断方法[J]. 南京航空航天大学学报(自然科学版), 2025, 57(3): 509-516.
- [3] 孙晓，李军，王磊. 轻量化 CNN 在齿轮箱故障诊断中的应用[J]. 制造业自动化, 2025, 47(3): 123-127.

作者简介：温霆超（2005.01-）男，汉族，本科在读，职称无，研究方向：数据计算。单位：上海工程技术大学。

课题项目：2025 年度校级大创项目：基于数据和模型融合的智能装备故障诊断研究（cxz2521001）