

# 模式分量能量特征赋能的齿轮传动系统 无监督故障溯源方法

毕志昊<sup>1</sup>, 姚锦涛<sup>1</sup>, 何清波<sup>1</sup>, 彭志科<sup>1,2</sup>

(1. 上海交通大学机械系统与振动全国重点实验室, 上海 200240; 2. 宁夏大学机械工程学院, 宁夏 银川 750021)

**摘要:** 齿轮传动系统是重大装备动力传递的关键, 然而传动系统易发生多样且难以预判的故障, 影响装备的运行安全与服役可靠性。针对齿轮传动系统在实际工程应用中的故障样本缺失及故障位置未知等问题, 提出了模式分量能量特征赋能的无监督故障溯源方法。该方法通过提取与传动元件故障机理相关、具有明确故障定位指向性的模式分量能量特征, 赋能无监督异常检测与可解释故障溯源。具体地, 使用所提特征构建无监督数据描述模型, 并引入归因解释方法量化各特征对异常判定结果的贡献度, 从而在无故障样本条件下, 实现齿轮传动系统故障元件的可解释溯源。在直升机主减速器行星台架与装甲车辆传动装置整机台架上, 分别进行单一故障与复合故障溯源验证。试验结果表明, 所提方法在无故障样本训练情况下, 可以实现准确的传动元件故障溯源。本方法为齿轮传动系统在缺乏故障数据情况下的故障溯源提供了有效解决途径, 具有工程推广价值与实际应用潜力。

**关键词:** 故障溯源; 齿轮传动系统; 特征提取; 归因解释

**中图分类号:** TH165+.3; TH132.41 **文献标志码:** A **文章编号:** 1004-4523(2025)06-1344-10

**DOI:** 10.16385/j.cnki.issn.1004-4523.2025.06.023

## Mode component energy feature-empowered unsupervised fault localization for gear transmission systems

BI Zhihao<sup>1</sup>, YAO Jintao<sup>1</sup>, HE Qingbo<sup>1</sup>, PENG Zhike<sup>1,2</sup>

(1. State Key Laboratory of Mechanical Systems and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China;

2. School of Mechanical Engineering, Ningxia University, Yinchuan 750021, China)

**Abstract:** Gear transmission systems are critical for power delivery in major equipment, yet they are susceptible to diverse and unpredictable faults, compromising operational safety and service reliability. In practical engineering applications, the absence of fault samples and unknown fault locations pose significant challenges. To address these issues, this paper proposes an unsupervised hypersphere-based fault localization (UHFL) method empowered by mode component energy features. This method extracts mode component energy features that are not only related to the fault mechanisms of transmission components but also possess clear fault localization interpretability. These features then enable both unsupervised anomaly detection and interpretable fault localization. Specifically, an unsupervised data description model is constructed using the proposed features. An attribution explanation method is introduced to quantify the contribution of each feature to the anomaly detection result, thereby achieving interpretable fault localization of gear transmission system components under conditions lacking fault samples. The proposed method is validated through single-fault and compound-fault localization experiments conducted on a helicopter main reducer planetary stage test bench and an armored vehicle transmission system test bench. Experimental results demonstrate that the proposed UHFL method can accurately localize faults in transmission components without requiring any fault sample training. This method offers an effective solution for fault localization in gear transmission systems under data-scarce conditions, showcasing valuable engineering promotion potential and application prospects.

**Keywords:** fault localization; gear transmission system; feature extraction; attribution analysis

齿轮传动系统作为重大装备的核心动力传递装置, 其可靠性直接关系到装备全寿命周期运行安

收稿日期: 2025-04-27; 修订日期: 2025-05-25

基金项目: 上海市市级科技重大专项资助课题; 基础研究资助项目(MKF20210013); 国家“两机”重大专项基础研究资助项目(J2019-IV-0018-0086)

全。尽管基于振动信号的异常检测技术已取得显著进展,但工程实践表明,异常检测仅为基础问题,更具工程价值的研究方向在于突破异常检测到故障溯源的壁垒,将检测结果精准定位至具体的故障元件,为运行维护提供决策依据。然而,重大装备齿轮传动系统的故障溯源面临两大关键难题:首先,由于系统固有的结构复杂性,多个传动元件间相互耦合且故障源呈现多样性,基于故障模拟试验数据的故障诊断方法难以应对所有潜在故障位置;其次,受制于装备运行安全的严苛要求,实际运行中可获取的故障样本极度稀缺,这使得依赖大量标签数据的监督学习方法在实际工程中应用严重受限,效果不佳。

在样本稀缺与标签缺失的情况下,弱监督学习框架下的故障溯源方法成为近年来的研究焦点。被测设备故障样本匮乏情况下的故障溯源技术发展主要围绕3个维度:(1)迁移学习方法<sup>[1]</sup>,通过跨域知识迁移机制,将源域积累的故障知识迁移至目标域,有效应对跨工况、跨设备以及复合故障的溯源需求;(2)数据增强方法<sup>[2]</sup>,依托生成对抗网络或物理模型,通过特征对抗学习与物理模型生成拟真样本,提升小样本条件下的模型分类能力;(3)归因解释方法<sup>[3]</sup>,通过构建特征贡献度量化模型,建立模型决策与特征的可解释关联。常用的基于代理模型的可解释性分析主要包括局部可解释模型<sup>[4]</sup>(local interpretable model-agnostic explanations, LIME)和夏普利加性解释模型<sup>[5]</sup>(SHapley additive exPlanations, SHAP),它们在无故障数据情况下展现出独特的优势。LIME通过构建局部代理模型,针对特定样本在特征空间邻域内生成扰动数据,进而得到样本各特征在被解释模

型中的近似结果。MEY等<sup>[6]</sup>将LIME框架应用于频谱分析,通过贡献度排序,成功识别出与传动系统失效相关的关键频率和阶次成分。SHAP方法则基于合作博弈理论,通过夏普利值量化每个特征对决策结果的边际贡献。BRITO等<sup>[7]</sup>在轴承故障诊断中使用SHAP分析振动信号时域与频域特征的贡献度,验证了该方法在无监督场景下的溯源可行性,并将溯源结果定位到具体的特征层面。HERWIG等<sup>[8]</sup>将SHAP方法嵌入到齿轮传动系统诊断方法中,将输入的时域振动信号变换到频域和时频域的解释域进行故障溯源分析。ZHONG等<sup>[9]</sup>进一步扩展SHAP在时频图故障溯源中的应用,为无监督域适应模型提供可视化的溯源途径。YU等<sup>[10]</sup>使用SHAP方法分析复合故障中时域和频域特征的贡献度,为多故障场景提供新的分析视角。通过代理模型构建的归因解释方法虽取得初步进展,但在处理传动系统多源故障时仍存在特征物理可解释性不足的缺陷。

当前无监督故障溯源方法的研究主要存在三方面局限性:第一,数据驱动方法在样本生成过程中过度依赖仿真数据或有限标签知识,难以保证数据的真实性,难以反映复合故障和未知故障场景;第二,基于特征归因的可解释性方法虽能提供特征层面的软解释,但难以定位到具体的传动元件;第三,现有研究多聚焦于故障类型识别层面,鲜有针对多级齿轮传动系统中的传动元件故障溯源研究。

面对上述挑战,受可解释人工智能中先验赋能与归因解释思想的启发,本研究通过构建具有物理可解释性的特征进行无监督故障溯源定位。如图1所示,通过构建与传动元件密切相关的模式分量能

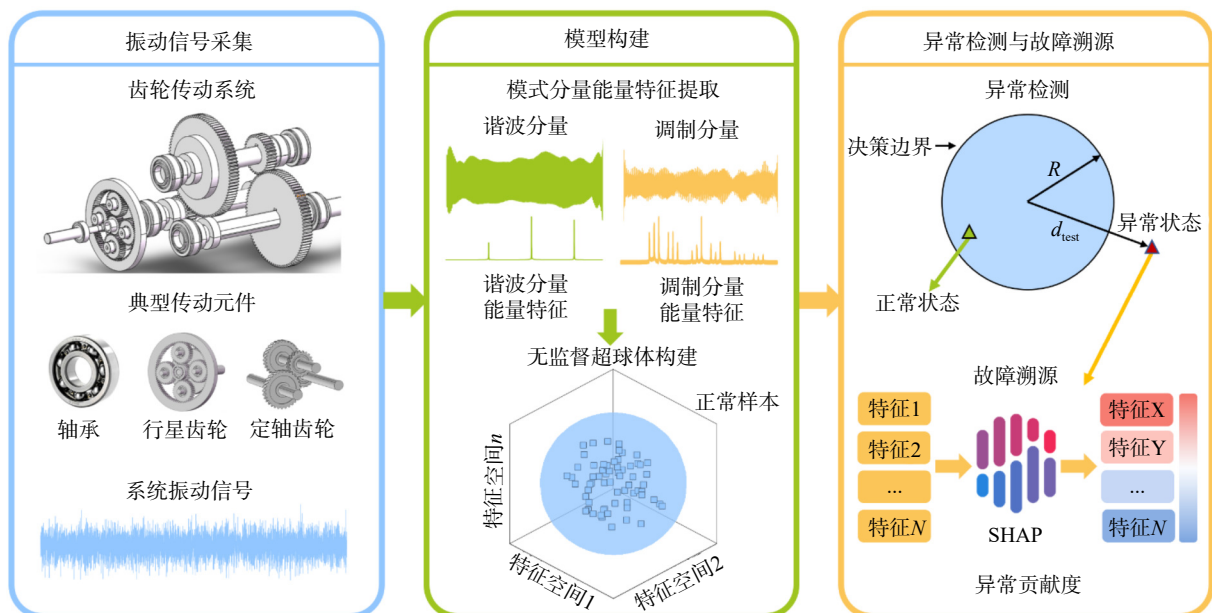


图1 可解释特征赋能的齿轮传动系统无监督故障溯源方法原理图

Fig. 1 Schematic of interpretable feature-enabled unsupervised fault localization method for gear transmission systems

量特征,建立特征与传动元件的显式关联,并结合归因解释方法,发展无需故障样本支撑的无监督溯源方法,最终形成可解释特征赋能的齿轮传动系统无监督故障溯源方法(unsupervised hypersphere-based fault localization, UHFL)。该方法旨在解决仅可获取正常状态数据条件下的故障溯源难题,实现无需依赖故障样本的传动元件故障追溯。相较于现有方法,UHFL方法的创新性体现在两个维度:首先,在特征构建层面,突破传统特征的物理解释壁垒,建立基于故障机理的可解释特征;其次,在方法架构层面,实现异常检测与故障溯源的有机融合,有效解决故障位置未知的故障元件定位难题。

为验证UHFL方法的工程适用性,本文设计了两个验证案例:首先在直升机主减速器行星架的单一故障上验证方法的故障溯源准确率;其次通过装甲车辆传动装置整机台架的复合故障,验证本方法在复合故障场景下的故障溯源能力。试验及对比结果展现了本文所提方法在真实工业场景中相比于主流无监督故障溯源方法的优势。

本文结构安排如下:第1节介绍可解释特征构建方法;第2节介绍基于归因解释的故障溯源方法;第3节详细阐述本文所提的无监督故障溯源方法;第4节为试验验证,分别在直升机主减速器行星架和装甲车辆传动装置台架上验证本方法在多故障单一故障和复合故障中的故障溯源能力;第5节为本文的研究结论。

## 1 可解释特征构建方法

齿轮传动系统通常工作于多变且严苛的环境中,系统内部结构复杂,包含大量传动元件,其振动信号中不同部件的特征频率成分彼此相邻、分布密集。本文首先提出基于信号多分量精确分解的模式分量能量特征构建方法,显著降低多源信号混叠对特征构建的影响。该特征构建方法充分融合了齿轮传动系统的先验结构信息,具备故障溯源可解释性,为无监督故障溯源提供了可靠的特征基础。

齿轮传动系统的振动信号可视为多个传动元件信号分量叠加而成。本文将系统振动信号 $S(t)$ 表示为源信号 $s_s(t)$ 与残差信号 $r(t)$ 的线性组合:

$$S(t) = \sum_{s=1}^{n_s} s_s(t) + r(t) = \sum_{s=1}^{n_s} \left\{ u_s(t) \cos \left[ 2\pi \int_0^t f_s(\tau) d\tau \right] + v_s(t) \sin \left[ 2\pi \int_0^t f_s(\tau) d\tau \right] \right\} + r(t) \quad (1)$$

式中, $u_s(t)$ 和 $v_s(t)$ 分别表示信号分量的幅值函数;

$n_s$ 为信号分量的总数量。

对信号分量的幅值函数和瞬时频率 $f_s(\tau)$ 进行傅里叶级数建模,并求解相应参数,实现信号分量的精确拟合,从而达到高精度的信号分解:

$$u_s(t) = u_0^{(s)} + \sum_{l=1}^L \left[ u_l^{(s)} \cos(2\pi l F_0 t) + \bar{u}_l^{(s)} \sin(2\pi l F_0 t) \right] \quad (2)$$

$$v_s(t) = v_0^{(s)} + \sum_{l=1}^L \left[ v_l^{(s)} \cos(2\pi l F_0 t) + \bar{v}_l^{(s)} \sin(2\pi l F_0 t) \right] \quad (3)$$

$$f_s(t) = f_c^{(s)} + \sum_{m=1}^M \left[ \beta_m^{(s)} \cos(2\pi m F_0 t) + \bar{\beta}_m^{(s)} \sin(2\pi m F_0 t) \right] \quad (4)$$

式中, $L$ 和 $M$ 表示傅里叶级数的阶数; $u_0^{(s)}, \dots, u_L^{(s)}, \bar{u}_1^{(s)}, \dots, \bar{u}_L^{(s)}, v_0^{(s)}, \dots, v_L^{(s)}$ 和 $\bar{v}_1^{(s)}, \dots, \bar{v}_L^{(s)}$ 为傅里叶系数; $f_c^{(s)}$ 为信号分量载波频率; $\beta_1^{(s)}, \dots, \beta_M^{(s)}$ 和 $\bar{\beta}_1^{(s)}, \dots, \bar{\beta}_M^{(s)}$ 表示冗余傅里叶模型的参数集合; $t = t_0, \dots, t_{T-1}$ 表示信号采样时刻,其中 $T$ 表示采样点数; $f_s$ 表示信号采样频率, $F_0 = f_s/(2T)$ 。利用本征啁啾分量分解法<sup>[11]</sup>对模型中的参数进行求解,从而实现传动系统各部件的模式分量提取。

对于齿轮传动系统中的轴承部件,提取轴承特征频率处的分量。轴承的典型特征频率包括外圈故障特征频率 $f_o$ 、内圈故障特征频率 $f_i$ 和滚动体故障特征频率 $f_b$ ,具体公式如下:

$$f_o = \frac{n \cos \phi_b (D - d)}{2D} f_r \quad (5)$$

$$f_i = \frac{n \cos \phi_b (D + d)}{2D} f_r \quad (6)$$

$$f_b = \frac{D^2 - d^2 \cos^2 \phi_b}{2Dd} f_r \quad (7)$$

式中, $f_r$ 为轴承转速; $n$ 为滚动体数量; $d$ 和 $D$ 分别表示滚动体直径和节圆直径; $\phi_b$ 为接触角。由于轴承故障通常会激发与轴承特征频率相关的基频及其高阶谐波成分,因此,构建轴承特征频率及其高阶谐波分量能量均值 $H_B$ ,其计算公式如下:

$$H_B = \frac{1}{n_{bf}} \sum_{k=1}^{n_{bf}} E_{bk} \quad (8)$$

式中, $E_{bk}$ 表示第 $k$ 阶特征频率处的包络信号能量; $n_{bf}$ 为考虑的最高特征频率阶数, $k = 1, 2, \dots, n_{bf}$ 。

对于每对啮合齿轮副,构建其啮合分量能量特征,定义为啮合分量及其高次谐波分量的能量的算术均值 $H_G$ :

$$H_G = \frac{1}{n_{mf}} \sum_{p=1}^{n_{mf}} E_{gp} \quad (9)$$

式中, $E_{gp}$ 表示 $p$ 阶啮合分量的能量; $n_{mf}$ 表示考虑的啮合分量最高阶数, $p = 1, 2, \dots, n_{mf}$ 。

构建调制分量能量特征,其定义为齿轮各阶啮合分量处的调制分量能量的算术均值 $H_M$ :

$$H_M = \frac{1}{n_{mf}} \sum_{p=1}^{n_{mf}} \frac{\sum_{q=1}^{n_{eb}} E_{bqp}}{E_{gp}} \quad (10)$$

式中,  $E_{bqp}$  表示  $p$  阶啮合分量在第  $q$  个调制边频处分量的能量;  $n_{eb}$  为最高调制边频数,  $q = 1, 2, \dots, n_{eb}$ 。

对于包含  $A$  个关键元件的齿轮传动系统, 构建模式分量能量特征集合  $\Omega = [CI^1, CI^2, \dots, CI^A]$ 。其中, 第  $a$  个传动元件所对应的模式分量能量特征集合表示为  $CI^a = [H_B^a, H_G^a, H_M^a]$ , 第  $i$  个样本的能量特征集合表示为  $x_i = [CI_i^1, CI_i^2, \dots, CI_i^A]$ 。

## 2 基于归因解释的故障溯源方法

本文所提出的无监督故障溯源方法, 首先利用构建的模式分量能量特征集合训练无监督数据描述模型, 对系统的状态进行判别。当测试样本为异常时, 进一步引入归因解释方法中的 SHAP 方法, 计算该样本中各特征对异常判定结果的贡献度。结合模式分量能量特征与齿轮传动系统结构之间的映射关系, 实现对异常状态的传动元件溯源定位。

无监督超球体数据描述方法旨在构建一个能够包络绝大多数正常样本的最小体积超球体。该方法仅依赖正常状态数据即可建立异常检测模型, 因而适用于实际工程中仅能获取有限健康数据的无标签场景。假设训练集  $X_{\text{train}} = \{x_i; i = 1, 2, \dots, N\}$  由  $N$  个样本组成, 超球中心为  $\mathbf{a}$ , 半径为  $R$ , 则构建最小超球体的目标函数表示为:

$$\begin{aligned} \min R^2 + \frac{1}{N} \sum_i (\varphi(x_i) - \mathbb{C})^T (\varphi(x_i) - \mathbb{C}) + \frac{1}{vN} \sum_i \zeta_i \\ \text{s.t. } (\varphi(x_i) - \mathbb{C})^T (\varphi(x_i) - \mathbb{C}) \leq R^2 + \zeta_i; \zeta_i \geq 0, \forall i \end{aligned} \quad (11)$$

式中,  $\zeta_i$  为松弛变量, 用于控制边界外样本, 增强模型的鲁棒性;  $\mathbb{C}$  为惩罚因子, 用于平衡超球体体积和误差;  $\varphi$  为隐式映射, 将输入样本从原始空间映射至高维特征空间。为求解该优化问题, 引入拉格朗日乘子  $\alpha_i \geq 0$  和  $\gamma_i \geq 0$ , 并结合 Karush-Kuhn-Tucker (KKT) 条件, 可得出:

$$\mathcal{L} = R^2 + \mathbb{C} \sum_i \zeta_i - \sum_i \gamma_i \zeta_i - \sum_i \alpha_i \{R^2 + \zeta_i - \|\varphi(x_i) - \mathbf{a}\|^2\} \quad (12)$$

求解上述问题时,  $\mathcal{L}$  对于  $R$ 、 $\mathbf{a}$ 、 $\zeta_i$  最小化, 对于  $\alpha_i$ 、 $\gamma_i$  最大化。  $K(\mathbf{x}_i \cdot \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \sigma^2)$  为高斯核函数, 表示隐含映射的内积,  $\sigma$  为核函数宽度系数。将上式的偏导数置零, 约束条件转换为:

$$\mathcal{L} = \sum_i \alpha_i K(\mathbf{x}_i \cdot \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j K(\mathbf{x}_i \cdot \mathbf{x}_j) \quad (13)$$

计算被测样本在超球体空间中到超球中心  $\mathbf{a}$  的距离平方  $R^2$  以进行正异常判断。无监督数据描述方

法通过下式对测试集中的被测样本  $\mathbf{x}_z$  进行异常检测:

$$\begin{aligned} HDD(\mathbf{x}_z) = \|\varphi(\mathbf{x}_z) - \mathbf{a}\|^2 = K(\mathbf{x}_z \cdot \mathbf{x}_z) - \\ 2 \sum_i \alpha_i K(\mathbf{x}_z \cdot \mathbf{x}_i) + \sum_{i,j} \alpha_i \alpha_j K(\mathbf{x}_i \cdot \mathbf{x}_j) \end{aligned} \quad (14)$$

齿轮传动系统的工作状态由  $HDD(\mathbf{x}_z)$  确定:

$$\text{工作状态} = \begin{cases} \text{正常}, & HDD(\mathbf{x}_z) \leq R^2 \\ \text{异常}, & HDD(\mathbf{x}_z) > R^2 \end{cases} \quad (15)$$

由于无监督数据描述方法引入了核函数将特征非线性映射至高维特征空间, 在增强模型的判别能力同时也使得模型预测结果与特征的对应关系变得复杂。开展故障溯源的关键在于明确模型在进行异常识别时依赖的主要特征。

基于特征贡献度排序的归因解释方法是一类典型的事后可解释性技术。此类方法针对模型对特定样本的预测输出, 通过量化各输入特征对结果的影响获得模型的决策依据。常用方法包括: 排列式特征重要性排序<sup>[12]</sup>、递归贡献度排序 (recursive contribution sorting, RCS)<sup>[13]</sup>、模型无关的 LIME 以及 SHAP 等。此类方法可辅助用户理解模型的决策逻辑, 并结合领域知识判断其输出结果的合理性, 从而提升模型在实际应用中的可用性与信任度。

SHAP 是一种广泛应用的加性特征归因方法, 其本质是将模型的输出结果表达为各输入特征的归因值之和。SHAP 方法融合博弈论与局部可解释性思想, 通过计算每个特征在所有可能特征子集中的夏普利值, 量化其对模型输出的影响。

引入一个用于局部解释的伴随模型  $g$ , 其为一个由二元特征组成的加性线性模型, 定义如下:

$$g(\mathbf{z}') = \psi_0 + \sum_{j=1}^J \psi_j z'_j \quad (16)$$

式中,  $J$  表示输入特征总数, 即全部用于检测的模式分量能量特征个数;  $\psi_0$  为解释模型的基线常数, 取训练样本预测值的均值;  $\psi_j$  为特征  $j$  的夏普利归因值。式 (16) 表示某个测试样本的模型预测结果等于其各输入特征的归因值之和加上基线值  $\psi_0$ 。  $z'_j \in \{0, 1\}$  表示特征是否被观测到:

$$z'_j \in \begin{cases} 0, & \text{第 } j \text{ 个特征不被观测} \\ 1, & \text{第 } j \text{ 个特征被观测} \end{cases} \quad (17)$$

夏普利值定义为特征子集的边际贡献, 具体定义如下:

$$\psi_j = \sum_{\mathbb{S} \subseteq \Omega \setminus \{j\}} \frac{|\mathbb{S}|!(J - |\mathbb{S}| - 1)!}{J!} [f(\mathbb{S} \cup \{j\}) - f(\mathbb{S})] \quad (18)$$

式中,  $\Omega$  表示构建的可解释特征集合,  $\Omega \setminus \{j\}$  表示不包含第  $j$  个特征的特征子集,  $\mathbb{S}$  为  $\Omega \setminus \{j\}$  的子集;  $f(\mathbb{S})$  表示初始模型的预测值。若测试集包含  $M$  个样本, 则第  $j$  个特征的全局重要性可由其在所有样本中归因值的绝对值均值表示:

$$G_j = \frac{1}{M} \sum_{m=1}^M |\psi_{jm}| \quad (19)$$

式中,  $\psi_{jm}$  为第  $j$  个特征在第  $m$  个测试样本的夏普利值。

通过量化与汇总测试样本中各个与关键部件密切相关的模式分量能量特征对模型异常检测结果的贡献度, 获得可解释特征的全局重要性排序, 并结合构建的可解释特征与齿轮传动系统的物理结构之间的关联关系, 在无需故障样本的情况下将异常状态有效溯源至特定传动元件。

### 3 提出的无监督故障溯源方法

本文提出的可解释特征赋能的齿轮传动系统无监督故障溯源方法原理图如图 1 所示。该方法首先针对齿轮传动系统的关键元件构建具有物理可解释性的模式分量能量特征, 此类特征能够为后续异常检测与故障溯源提供基础支持。

在对正常状态信号和被测信号分别提取可解释特征之后, 基于正常状态的特征训练无监督数据描述模型。该模型通过在特征空间中构建包络正常样本的最小超球体, 从而划定系统的正常状态的决策边界。对于输入样本, 若其在特征空间中落入超球体内部, 则判定为正常状态; 若其位于超球体之外, 则视为异常样本。

针对检测出的异常样本, 引入归因解释的夏普利加性解释模型, 计算每个特征的归因值, 以量化其对异常判定结果的边际贡献, 并据此进行特征排序。由于所提特征具备明确的物理含义, 贡献度最高的特征可直接映射至对应的传动元件, 从而实现故障部位的精准溯源。

最终, 结合归因解释结果与特征所对应的系统结构信息, 可将异常状态精确定位至具体的齿轮、轴承等关键元件。本文提出的方法主要用于无故障数据训练场景下的故障溯源任务, 其算法伪代码如表 1 所示, 算法流程图如图 2 所示。

### 4 试验验证

本节通过两个试验案例验证所提出的可解释特征赋能的无监督故障溯源方法的有效性与其适用性。针对直升机主减速器行星级台架的单一故障试验, 验证本方法在多故障源情况下的单一故障溯源能力; 进一步, 在装甲车辆传动装置台架的复合故障场景中, 验证其在多源复合故障情况下的故障溯源性能。

表 1 提出的无监督故障溯源方法伪代码

Tab. 1 Pseudocode of the proposed UHFL method

算法 1: 提出的无监督故障溯源方法算法

输入: 正常状态信号和被测信号

特征提取:

获取与故障相关的模式分量并构建可解释的模式分量能量特征, 得到训练集合  $X_{\text{train}}$  和测试样本  $x_z$

无监督数据描述模型构建:

使用正常状态特征集合  $X_{\text{train}}$  构建数据描述超球体, 将决策边界  $R^2$  作为异常检测阈值

故障溯源:

For 被测样本特征  $x_z$

计算  $x_z$  到超球中心的距离  $HDD(x_z)$

If  $HDD(x_z) \leq R^2$

系统处于正常状态

Else

系统处于异常状态

构建伴随模型  $g(z')$

For 异常样本的全部特征  $J$

计算特征的夏普利归因值  $\psi_j$

排序归因值  $\psi_j$

依据模式分量能量特征获得溯源结果

For 循环结束

End If

For 循环结束

输出: 故障溯源结果

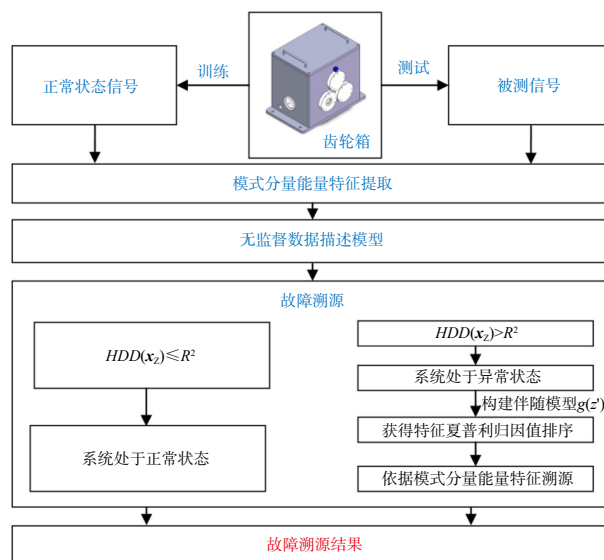


图 2 提出的无监督故障溯源方法流程图

Fig. 2 Flowchart of the proposed UHFL method

#### 4.1 直升机主减速器行星级台架单一故障溯源

为验证所提出的 UHFL 方法的准确性, 在转速为 2400 r/min、负载为 1200 N·m 的工况下, 于行星级试验台架上开展试验测试。试验设置了两类典型故障位置, 包括行星齿轮和太阳轮, 并设计三种故障类型: 行星轮点蚀、行星轮裂纹及太阳轮裂纹, 测试台架与故障件如图 3 所示。在该系统中, 太阳轮为驱动端、齿数为 34, 齿圈固定齿数为 106, 4 个行星轮齿

数为 36, 在齿圈的约束下围绕太阳轮公转, 并最终通过行星架输出扭矩。信号采样频率为 20 kHz, 每类振动信号划分为 238 个样本, 图 4 展示了不同状态下的振动波形与频谱, 表 2 给出了对应的特征频率。

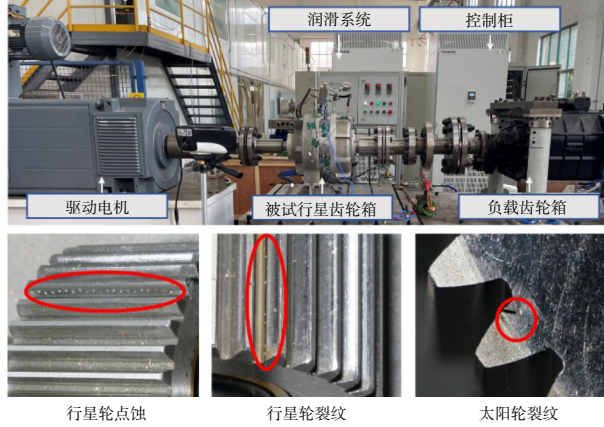


图 3 主减速器行星轮试验台架与故障件详情

Fig. 3 Overview of main reducer planetary gear test system and fault components

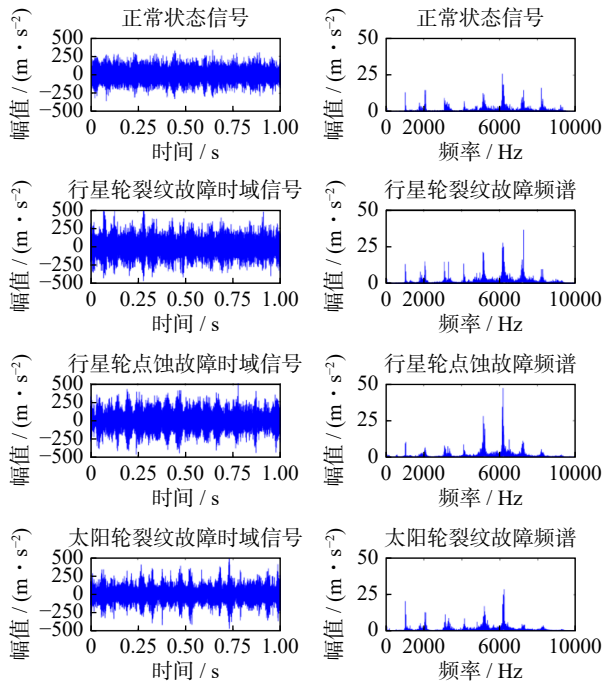


图 4 原始振动信号波形(左)及频谱(右)

Fig. 4 The original vibration waveform (left) and the frequency spectrum (right)

表 2 主减速器行星轮试验台架故障特征频率

Tab. 2 Characteristic frequency of the main reducer planetary gear test bench

$f_{\text{shaft}}/\text{Hz}$	$f_{\text{mesh}}/\text{Hz}$	$f_{\text{planet}}/\text{Hz}$	$f_{\text{sun}}/\text{Hz}$
40.00	1029.71	28.60	30.29

无监督故障溯源方法分为异常检测和故障溯源两个阶段。首先, 随机选取一半正常样本作为训练集, 用于构建异常检测模型, 其余正常样本与所有故

障样本构成测试集。如图 5 所示, UHFL 能够对测试集中所有样本实现准确的状态识别。

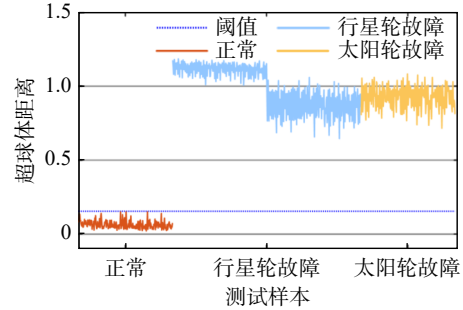


图 5 UHFL 的主减速器行星轮试验台架异常检测结果

Fig. 5 The main reducer planetary gearbox anomaly detection result of the UHFL method

在此基础上, 进一步对异常样本进行故障溯源, 识别对异常判定影响最大的关键特征, 并结合特征与元件的结构映射关系, 完成故障部位的定位。试验结果表明, 所提方法可成功将异常溯源至实际故障的传动元件。尽管如表 2 所示的行星轮与太阳轮特征频率接近, 但通过信号分量分解与能量特征构建, 模型能够有效区分二者。最终, 基于 SHAP 方法的特征贡献度分析, 实现了对具体故障部件的准确识别与定位。

在故障溯源任务中, 本文将所提出的 UHFL 方法与主流无监督故障溯源方法进行对比, 所有方法均将模式分量能量特征作为输入。本案例选取了 4 种常用的无监督故障溯源方法进行对比, 分别为: 主成分分析+平方误差预测 (PCA+SPE)<sup>[14]</sup>、高斯混合模型 (GMM)<sup>[15]</sup>、多类失衡学习 (multiclass imbalanced learning, MIL)<sup>[16]</sup>、无监督深度迁移学习 (unsupervised deep transfer learning, UDTL)<sup>[17]</sup>。为了量化故障溯源性能, 采用准确率 *Accuracy* 作为评估指标, 其计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

式中, *TP*、*TN*、*FP* 和 *FN* 分别为真正例、真负例、假正例和假负例的数量。

UHFL 与 PCA 方法在流程上具有相似性, 均采用“异常检测+特征溯源”的两阶段策略: 首先识别异常样本, 再基于特征贡献度完成故障部位定位。对比结果表明, PCA+SPE 方法在故障定位的稳定性和一致性方面表现良好, 但整体性能仍不及 UHFL 方法。与之相比, 聚类方法在溯源策略上存在根本差异, 其直接对三类故障样本进行聚类划分, 虽能区分类别间差异, 但缺乏对具体故障类型的识别能力, 且在无标签样本训练条件下, 难以实现准确的部件级定位。如表 3 所示, UHFL 方法在识别精度上明显优于其他无监督方法, 分类准确率为 99.86%, 较主流方法的最优结果提升 2.02%。综上, 本文所提的

UHFL 方法在行星齿轮箱内部故障溯源方面展现出显著优势。

表 3 主减速器行星级试验台架故障溯源准确率

Tab. 3 Fault localization results for the main reducer planetary gear test bench

方法	准确率/%
PCA+SPE	70.73
GMM	81.00
MIL	86.19
UDTL	97.84
UHFL	99.86

4.2 装甲车辆传动装置整机台架复合故障溯源

装甲车辆传动装置是动力传输系统的核心部件,负责将发动机输出的转速与扭矩传递至驱动轮。从结构上,传动装置主要由定轴传动机构与行星变速机构构成。其中的定轴传动机构负责将动力分配至前传动、风扇传动、变矩器等子系统;行星变速机构则承担换挡与变速功能。在严苛的真实服役环境下,传动装置易发生齿轮和轴承等关键部件故障,且其故障位置往往具有不确定性,甚至存在复合故障情形。由于传动装置结构紧凑、集成度高,故障定位难度大,因此,若能在无需拆解的情况下实现传动装置的故障溯源,将具有重要的工程价值。为此,本节将使用无监督故障溯源方法实现装甲车辆传动系统的异常检测与复合故障定位。

试验台架基于某型传动装置,其结构如图 6 所示。通过在定轴传动机构与行星变速机构两个关键子系统中植入典型故障件,构建覆盖常见故障类型及复合故障情形的试验场景。试验共设计了 11 种典型故障件,分布于 2 个子系统的 5 个故障频发位置,并组合成 3 套具有代表性的复合故障组合。故障类型涵盖齿轮裂纹、剥落,以及轴承内圈、外圈损伤与保持架断裂等形式,3 套故障组合分别由 3 件、4 件与 4 件故障传动元件组成。故障件及植入方案如图 7 所示,故障特征阶次如表 4 所示。

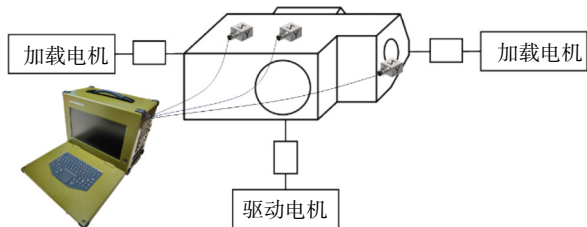


图 6 传动装置整机试验台架简图

Fig. 6 Sketch of the test bench for transmission system

关于故障组合的设计目的与关注重点的说明如下:故障组合 1,模拟仅定轴传动机构发生复合故障的情形;故障组合 2,模拟定轴与行星机构同时发生故障的情形;故障组合 3,模拟两个机构同时发生齿

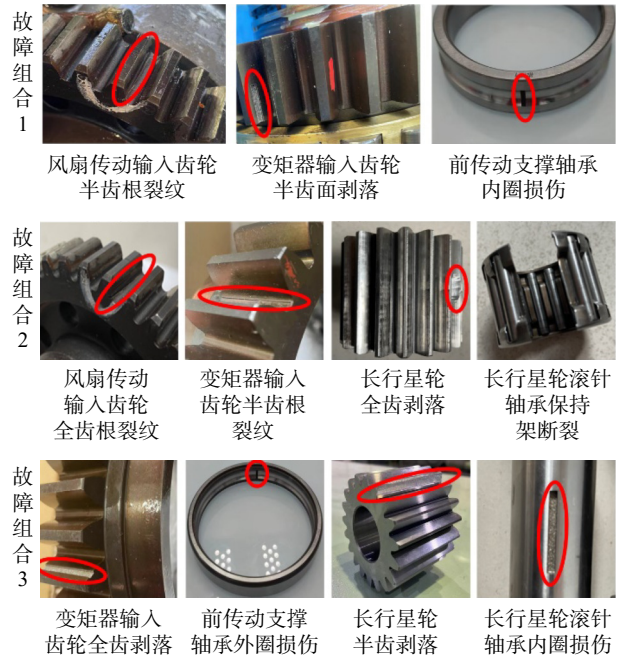


图 7 传动装置故障件组合详情

Fig. 7 Details of faulty component combinations for the transmission system

表 4 传动装置各部件特征阶次

Tab. 4 Components characteristic orders of the transmission system

故障特征阶次	定轴传动机构故障件			行星变速机构故障件	
	风扇传动输入齿轮	变矩器输入齿轮	前传动支撑轴承	长行星齿轮	长行星轮滚针轴承
啮合	59.0000	53.0350	-	89.1966	-
调制	1.5946	1.5599	-	3.0106	-
内圈	-	-	6.8322	-	24.0852
外圈	-	-	5.1678	-	18.0641
保持架	-	-	0.4306	-	1.7204

轮和轴承故障的情形。

为保证数据的一致性,在正常状态及 3 种故障组合状态下,使系统处于前进 4 挡、输入转速 2300 r/min、单边输出负载为 1036 N·m 的工况下采集振动加速度信号,信号波形如图 8 所示。

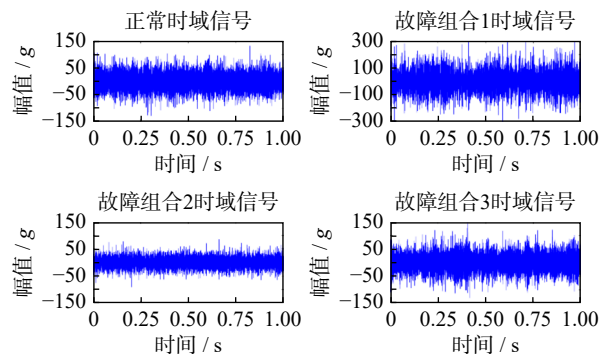


图 8 传动装置不同状态的振动信号

Fig. 8 The vibration signals of the transmission system under different conditions

所提的UHFL方法包含异常检测与故障定位两个阶段。首先,仅基于正常样本构建无监督数据描述模型,实现系统状态的异常识别。随后,将检测出的异常样本输入故障溯源流程,依据特征与位置的映射关系,实现具体故障部件溯源。提取各故障部件的模式分量能量特征,作为异常检测与溯源分析的输入。使用正常状态下提取的特征组合训练数据描述模型,其决策边界作为异常检测的判定阈值。如图9所示,本方法可实现对系统运行状态的准确识别。

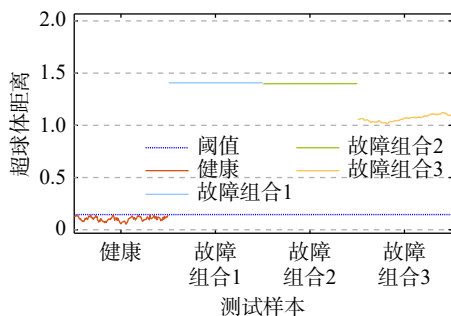


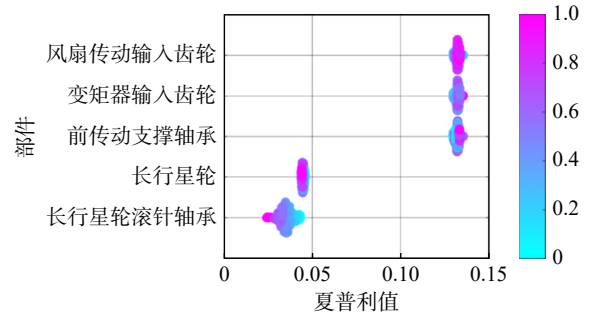
图9 UHFL方法的传动装置试验台架异常检测结果

Fig. 9 The transmission system anomaly detection result of the UHFL method

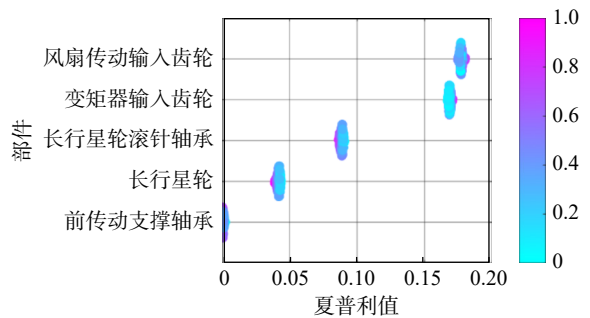
进一步,针对传动装置的5个潜在故障位置,均可通过各自构建的模式分量能量特征实现有效定位。尽管齿轮部件的啮合阶次与调制阶次较为接近,但得益于信号多分量分解方法与能量特征的差异性构建,所提特征可有效定位故障部位。

图10展示了三组故障组合的故障溯源结果,图中纵坐标为各部件位置对应的模式分量能量特征,横坐标为其综合夏普利值,颜色表示该特征对异常预测结果的影响强度。从图10(a)可见,风扇传动输入齿轮、变矩器输入齿轮及前传动支撑轴承的特征贡献度明显高于其他部件,准确对应故障组合1的实际故障位置。图10(b)中风扇输入齿轮、变矩器输入齿轮、长行星轮及滚针轴承的贡献值较高,未植入故障的前传动支撑轴承贡献度显著较低。图10(c)显示,在故障组合3中,长行星轮、滚针轴承及变矩器输入齿轮的特征贡献度占据主导,前传动支撑轴承也呈现出明显的贡献值,而风扇传动输入齿轮几乎无贡献。综上,即使在复合故障条件下,所构建的基于模式分量能量特征的无监督故障溯源方法,仍可实现对故障部件的准确识别,体现出良好的解释性与鲁棒性。

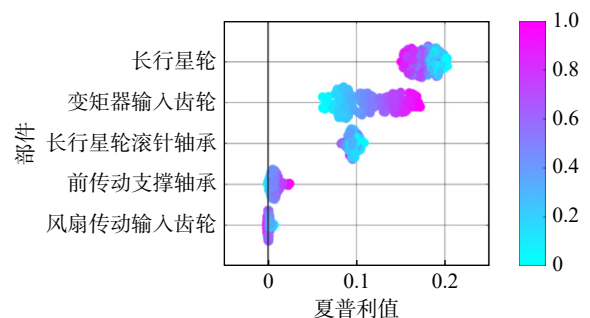
在故障溯源对比试验中,将本文所提出的UHFL方法与主流无监督故障溯源方法进行对比。所有方法均使用相同的模式分量能量特征作为输入。



(a) 故障组合1  
(a) Fault combination 1



(b) 故障组合2  
(b) Fault combination 2



(c) 故障组合3  
(c) Fault combination 3

图10 所提方法传动装置故障溯源结果

Fig. 10 Transmission system fault localization results of the proposed method

如图11所示,以GMM和K-Means为代表的聚类方法虽可将样本划分为4类,对应传动装置的4种状态,但本质上仅反映数据分布特征,缺乏对具体故障部件的识别能力。将UDFL与PCA+SPE和递归贡献度排序方法(RCS)<sup>[13]</sup>进行对比,这类方法与UHFL方法均采用“异常检测+特征溯源”的两阶段框架。以故障组合3为例,对比方法的溯源效果如图12所示,PCA+SPE方法和递归贡献度排序方法均能准确指示变矩器输入齿轮和长行星轮的故障。但在特征排序的区分能力上仍存在不足,与UHFL方法在故障组合3的结果图10(c)对比可以得到,现有无监督故障溯源方法在排序为第3~第5的特征之间区分度较低,限制了故障溯源精度。复合故障溯源准确率如表5所示。对比结果表明,UHFL方法在多部件复合故障场景下的故障

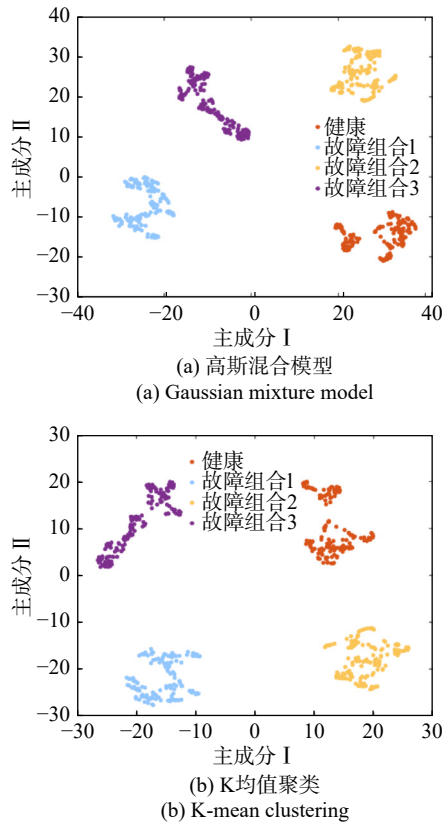


图 11 聚类方法的传动装置故障溯源结果对比

Fig. 11 Comparison of transmission system fault localization results of clustering methods

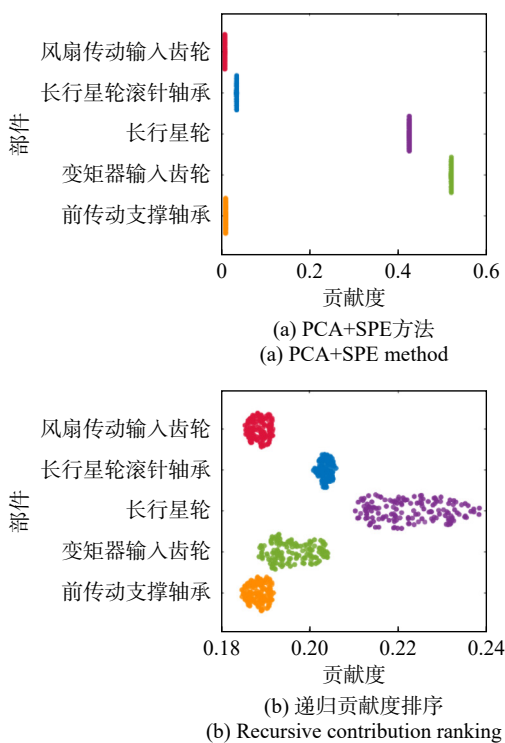


图 12 贡献度排序方法的故障组合 3 溯源结果对比

Fig. 12 Comparison of fault combination 3 localization results of contribution ranking methods

溯源准确率为 99.39%，具有更高的定位准确性与溯源能力。

表 5 传动装置试验台架故障溯源准确率

Tab. 5 Fault localization results for the transmission system

test bench	
方法	准确率%
PCA+SPE	69.44
RCS	92.78
<b>UHFL</b>	<b>99.39</b>

## 5 结论

针对齿轮传动系统在工程实践中面临的异常样本匮乏与无故障数据情况下故障溯源等难题，本文提出了一种基于可解释特征赋能的无监督故障溯源方法。该方法通过提取具有故障定位指向性的模式分量能量特征，在仅依赖正常数据进行训练的情况下，赋能无监督异常检测与基于特征归因解释的故障溯源方法，实现了传动系统故障的准确识别与传动故障元件定位。在直升机主减速器行星台架与装甲车辆传动装置整机台架上开展了单一故障和复合故障溯源的试验验证。结果表明：所提的 UHFL 方法在直升机台架上的多故障源单一故障溯源准确率为 99.86%，显著优于主流无监督方法；在装甲车辆传动装置台架上的复合故障溯源中，UHFL 亦能在无故障样本参与训练的条件下，准确定位多部件的复合故障，溯源准确率为 99.39%。综上，本文提出的方法为重大装备齿轮传动系统在无故障数据情况下的异常检测与故障溯源提供了有效解决途径，具有出色的工程推广价值与应用潜力。

## 参考文献:

- [1] LI W, HUANG R Y, LI J P, et al. A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: theories, applications and challenges[J]. *Mechanical Systems and Signal Processing*, 2022, 167: 108487.
- [2] 向家伟. 机械故障数值模拟驱动的生成式对抗网络及智能诊断原理[J]. *中国科学: 技术科学*, 2021, 51(3): 341-355.  
XIANG J W, Numerical simulation driving generative adversarial networks in association with the artificial intelligence diagnostic principle to detect mechanical faults [J]. *Scientia Sinica Technologica*, 2021, 51(3): 341-355.
- [3] 严如强, 周峥, 杨远贵, 等. 可解释人工智能在工业智能诊断中的挑战和机遇: 归因解释[J]. *机械工程学报*, 2024, 60(12): 21-40.  
YAN R Q, ZHOU Z, YANG Y G, et al. Challenges and opportunities of XAI in industrial intelligent diagnosis: attribution interpretation[J]. *Journal of Mechanical Engineering*,

- 2024, 60(12): 21-40.
- [4] SANAKKAYALA D C, VARADARAJAN V, KUMAR N. Explainable AI for bearing fault prognosis using deep learning techniques[J]. *Micromachines*, 2022, 13(9): 1471.
- [5] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. ACM, 2017: 4768-4777.
- [6] MEY O, NEUFELD D. Explainable AI algorithms for vibration data-based fault detection: use case-adapted methods and critical evaluation[J]. *Sensors*, 2022, 22(23): 9037.
- [7] BRITO L C, SUSTO G A, BRITO J N, et al. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery[J]. *Mechanical Systems and Signal Processing*, 2022, 163: 108105.
- [8] HERWIG N, BORGHESANI P. Explaining deep neural networks processing raw diagnostic signals[J]. *Mechanical Systems and Signal Processing*, 2023, 200: 110584.
- [9] ZHONG J H, LIN C, GAO Y, et al. Fault diagnosis of rolling bearings under variable conditions based on unsupervised domain adaptation method[J]. *Mechanical Systems and Signal Processing*, 2024, 215: 111430.
- [10] YU C C, LI M X, WU Z N, et al. Feature selection and interpretability analysis of compound faults in rolling bearings based on the causal feature weighted network[J]. *Measurement Science and Technology*, 2024, 35(8): 086201.
- [11] CHEN S, PENG Z, YANG Y, et al. Intrinsic chirp component decomposition by using Fourier Series representation[J]. *Signal Processing*, 2017, 137: 319-327.
- [12] ALTMANN A, TOLOŞI L, SANDER O, et al. Permutation importance: a corrected feature importance measure[J]. *Bioinformatics*, 2010, 26(10): 1340-1347.
- [13] BI Z H, YANG Y, DU M G, et al. Unsupervised hypersphere description approach for detecting and localizing anomalies in drivetrain with normal data[J]. *Measurement*, 2024, 228: 114349.
- [14] HE Q B, YAN R Q, KONG F R, et al. Machine condition monitoring using principal component representations[J]. *Mechanical Systems and Signal Processing*, 2009, 23(2): 446-466.
- [15] MALIUK A S, PROSVIRIN A E, AHMAD Z, et al. Novel bearing fault diagnosis using gaussian mixture model-based fault band selection[J]. *Sensors*, 2021, 21(19): 6579.
- [16] LI Q M, SONG Y J, ZHANG J, et al. Multiclass imbalanced learning with one-versus-one decomposition and spectral clustering[J]. *Expert Systems with Applications*, 2020, 147: 113152.
- [17] ZHAO Z B, ZHANG Q Y, YU X L, et al. Applications of unsupervised deep transfer learning to intelligent fault diagnosis: a survey and comparative study[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 3525828.
- 
- 第一作者:毕志昊(1997—),男,博士研究生。  
E-mail: bzhjh64@163.com
- 通信作者:何清波(1980—),男,博士,教授。  
E-mail: qbhe@sjtu.edu.cn