

基于多通道融合的齿轮箱故障诊断方法

冯志刚, 王 颖, 王 宇

(沈阳航空航天大学 自动化学院, 沈阳 110136)

摘 要: 针对单通道的振动信号在全面表征齿轮箱故障特征方面存在局限性的问题, 提出了一种将多通道的振动信号变换为时频图后经过加权融合成 RGB 图像, 再输入到诊断模型中的数据级融合方法 MFTDCN。该方法能有效整合多个通道的信息, 获得比单个振动信号更丰富的特征。首先, 对多通道振动信号进行主成分分析(PCA)处理, 降维成三个分量, 用短时傅里叶变换(STFT)将 3 个分量变换为 3 个时频图像; 然后, 用 Hoyer 稀疏度对余弦相似度进行改进, 将其应用于基于稀疏性和相似度的权重计算, 利用得到的权重对三个时频图像进行加权处理, 从而融合成 RGB 图像, 作为诊断模型的输入。诊断模型部分为可变形卷积(DCN)与 Transformer 相结合, 提取特征后进行分类。在东南大学的齿轮箱数据集和北京工业大学与北京交通大学联合发布的数据集上进行了相关实验验证, 结果证实了该方法的有效性和可靠性。

关键词: 故障诊断; 数据级融合; 短时傅里叶变换; 余弦相似度; 可变形卷积; Transformer

中图分类号: TH17; TP277 **文献标志码:** A **文章编号:** 1671-5497(2026)07-1811-14

DOI: 10.13229/j.cnki.jdxbgxb.20241377

Gear box fault diagnosis method based on multi-channel fusion

FENG Zhi-gang, WANG Ying, WANG Yu

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

Abstract: To address the limitations of single-channel vibration signals in comprehensively characterizing gearbox fault features, this paper proposes a data-level fusion method called MFTDCN is proposed. This method effectively integrates information from multiple channels to obtain richer features compared to single vibration signals. The approach begins by applying Principal Component Analysis (PCA) to multi-channel vibration signals to reduce them into three components. These components are then transformed into three time-frequency images using Short-Time Fourier Transform (STFT). An improved cosine similarity measure, based on Hoyer sparsity, is employed to calculate weights by combining sparsity and similarity metrics. These weights are used to perform weighted fusion of the three time-frequency images, resulting in a single RGB image, which is then used as input for the diagnostic model. The diagnostic model combines Deformable Convolution Networks (DCN) with Transformer architectures to extract features and classify faults. Experimental results on the gearbox data set of Southeast University and the data set jointly published by Beijing University of Technology and Beijing Jiaotong University confirm the effectiveness and reliability of the proposed method.

收稿日期: 2024-12-30.

基金项目: 国家自然科学基金青年基金项目(51605309).

作者简介: 冯志刚(1980-), 男, 教授, 博士. 研究方向: 故障诊断. E-mail: fzg1023@yeah.net

Key words: fault diagnosis; data-level fusion; short-time Fourier transform; cosine similarity; deformable convolution network; Transformer

0 引言

齿轮箱是一个由齿轮、轴承和传动轴组成的一个复杂系统,在运动转换和动力传递中发挥着关键作用^[1],且被广泛应用于汽车、货船、发电厂、机械制造及航空航天等众多领域^[2]。齿轮箱在高速重载的条件下运行,极易发生故障,可能引发安全事故,造成经济损失,甚至是人员伤亡。因此,能在早期阶段对齿轮箱的潜在故障进行准确、高效的诊断和监测,具有至关重要的意义^[3,4]。

现有的故障诊断工作主要是基于模型的方法和数据驱动的方法^[5]。然而,基于模型的传统故障诊断方法依赖于人工特征提取,并且需要专业知识支撑,过程既耗时又费力^[6]。近年来,随着技术的不断进步,数据驱动方法已成为研究热点,包括支持向量机(SVM)、随机森林、人工神经网络、决策树、K-最近邻(KNN)等^[7,8]。在一些研究中,将机器学习与信号处理技术相结合的方法较为常见,可以有效提升模型的诊断性能^[9]。Goyal等^[10]对信号进行希尔伯特变换去噪,PCA降维处理后,采用顺序浮动前向选择(SFFS)进行特征选择,并将选定的特征应用于SVM和人工神经网络(ANN)进行故障诊断;Zhang等^[11]提出了一种基于增强数据独立性的朴素贝叶斯(NB)和选择性支持向量机(SSVM)的轴承故障诊断方法。尽管机器学习技术在不断发展,但在处理复杂数据和多故障类别的机械传动部件时仍存在局限性。例如,BP算法和SVM易受样本数量及多故障、大数据等因素的影响,导致故障诊断的准确率降低^[12,13]。

深度学习(DL)提供了一个端到端的诊断方式,能自动从大量数据中提取深层次特征,并捕捉数据间的复杂相关性^[14]。因此,深度学习在一定程度上解决了传统机器学习的局限性。许多深度学习的诊断方法都将原始信号转换为二维图像作为模型的输入。Zhu等^[15]将振动信号转换为对称点图之后,再输入到卷积神经网络(CNN)中,对图像进行自适应特征提取,并验证了该方法的有效性;Wang等^[16]将振动信号转换为二维时频图

后作为改进的深度残差网络(IDRN)模型的输入,解决了数据不平衡的问题;Li等^[17]提出了一种将拉格姆角场(GAF)和CNN结合的故障诊断方法,使平均诊断准确率达到98.4%;Ding等^[18]提出了一种基于小波包能量(WPE)图像和CNN的能量波动多尺度特征挖掘方法,用于主轴轴承的故障诊断。本文将采用RGB图像作为深度学习模型输入,应用于齿轮箱的故障诊断。

然而,基于深度学习的齿轮箱故障诊断方法通常以单通道振动信号作为输入。由于齿轮箱的复杂性,仅用单通道振动信号难以全面反映故障信息。因此,为了获得更加全面的故障信息,本文设计一种多通道振动信号的数据级融合方式对齿轮箱进行故障诊断。该方法通过数据级融合方式,能够充分整合多个通道的互补信息,显著提升了故障诊断的准确性与可靠性,尤其适用于信号特征模糊、故障信息冗余的复杂场景。并且,融合策略精准高效,通过基于稀疏度改进的余弦相似度权重计算机制,可自适应地突出关键通道特征,避免冗余信息干扰。总的来说,该方法有效克服了单通道振动信号信息不足、难以充分表征齿轮箱复杂故障特征的问题。

1 理论基础

1.1 主成分分析(PCA)

主成分分析(PCA)是一种常用的线性降维方法,主要用于高维数据中的特征提取和数据压缩。PCA在尽可能保留原有信息的基础上提取互不相关的主成分,通过各主成分的贡献率选出重要特征作为模型的输入,所以PCA可以减少特征数量,达到降维的效果。

PCA过程和特征提取具体步骤如下:

(1)设原始数据为 $X_0 \in R_{m \times n}$,其中, m 为样本个数; n 为特征变量的个数。

(2)将原始数据标准化,均值为0,方差为1,计算公式如下:

$$X_s = \frac{X_0 - \text{mean}(X_0)}{\text{std}(X_0)} \quad (1)$$

式中: X_s 为标准化矩阵; $\text{mean}(\cdot)$ 为 $\cdot$ 的均值; $\text{std}(\cdot)$ 为 $\cdot$ 的标准差。

(3) 计算数据的协方差矩阵, 计算公式如下:

$$C = \frac{X_s^T X_s}{n-1} = U \Lambda U^T \quad (2)$$

式中: $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, $\lambda_1, \lambda_2, \dots, \lambda_n$ 为特征值, 且 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$; $U = [p_1, p_2, \dots, p_n]$ 是由特征值对应的标准正交特征向量组成的矩阵。

(4) 计算前 k 个主成分的累计方差贡献 (CCV), 通过特征值的大小, 选择前 k 个特征值对应的特征向量, 作为数据降维的主成分。

$$\eta_k = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i} \quad (3)$$

式中: η_k 为累计方差贡献率。

累计方差贡献率用来估计新提取的分量对原始数据的信息保留程度, 一般要求大于 85%。

(5) 将选择的 k 个特征值对应的特征向量构成变换矩阵 P :

$$P = [p_1, p_2, \dots, p_k], k < n \quad (4)$$

式中: P 为 n 行 k 列的矩阵。

(6) 将数据通过矩阵 P 转换为低维主成分矩阵。计算公式如下:

$$X = X_s P \quad (5)$$

式中: X 为降维后的数据矩阵, 维度为 $m \times k$ 。

通过以上过程的 PCA 特征提取, 将原始数据映射到新的坐标系中, 并选择方差最大的方向作为主成分, 实现有效降维。这种处理方法在保留振动信号主要故障信息的同时, 去除了冗余数据, 使生成的 RGB 图像能够更清晰地体现关键故障特征, 提高模型识别的准确性。

1.2 可变形卷积网络 (DCN)

可变形卷积通过引入可学习的空间偏移 (offset), 可以动态调整采样位置, 使其能更灵活地捕捉特征, 相对于传统卷积神经网络 (CNN), 提高了处理复杂特征时的表现能力。

可变形卷积在传统采样位置的基础上增加了动态偏移量 Δr , 偏移量可根据输入特征图动态调整, 从而使网络自动选择最合适的特征采样位置。图 1 所示为可变形卷积示意图。

在可变形卷积中, 定义 $R = \{(a, b) | a, b \in \mathbb{N}^*, a \leq m, b \leq h\}$ 为输入特征图上所有采样点的索引集合, 对标准卷积进行变形, 扩大在特征图上的采样范围, 则对应新的采样索引点为 $p = p_0 + r + \Delta r$, 所以, 输出特征图在位置 p_0 的值表示为:

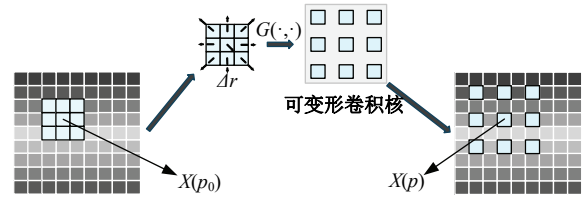


图 1 可变形卷积示意图

Fig. 1 Deformable convolution diagram

$$y(p_0) = \sum_{r \in R} \tilde{w}(r) X(p_0 + r + \Delta r) \quad (6)$$

Δr 是一个浮点数, 使用双线性插值对其进行采样。双线性插值公式如下:

$$X(p) = \sum_{q, p \in R} G(q, p) X(q) \quad (7)$$

$$G(q, p) = g(q_x, p_x) g(q_y, p_y) \quad (8)$$

$$g(a, b) = \max(0, 1 - |a - b|) \quad (9)$$

式中: $X(\cdot)$ 为双线性插值法计算采样点 p 处对应的采样值; q 表示特征图中有可能的整数空间位置; $G(\cdot, \cdot)$ 表示二维双线性插值核, 可分解为两个一维线性插值 $G(\cdot, \cdot)$ 。

1.3 Transformer 网络

Transformer 是一种基于注意力机制的深度学习网络构架, 依赖自注意力机制来建模序列中元素之间的全局依赖关系, 具有高度的并行化能力和灵活的结构表达能力。Transformer 网络主要由输入层、编码器和解码器构成。图 2 所示为 Transformer 结构图。

1.3.1 输入层

输入层的核心为位置编码模块, 由于自注意力机制会降低学习时序特征位置的能力, 因此, 需要对数据的位置进行编码来帮助模型捕捉时序关系, 位置编码器计算过程如下:

$$\begin{cases} PE_{(p, 2i)} = \sin\left(\frac{p}{10000^{2i/d_{\text{model}}}}\right) \\ PE_{(p, 2i+1)} = \cos\left(\frac{p}{10000^{2i/d_{\text{model}}}}\right) \end{cases} \quad (10)$$

式中: p 为序列中的位置索引 (从 0 开始); i 为对应的特征维度; $2i$ 和 $2i+1$ 分别为偶数和奇数维度; d_{model} 为位置编码的维度, 即输入特征的维数。

1.3.2 编码器

编码器主要由多头自注意力机制 (Multi-head self-attention)、前馈网络 (Feed forward network) 和残差连接 (Residual connection) 与层归一化 (Layer normalization) 组成, 其作用是将输入的特征序列编码为中间向量。

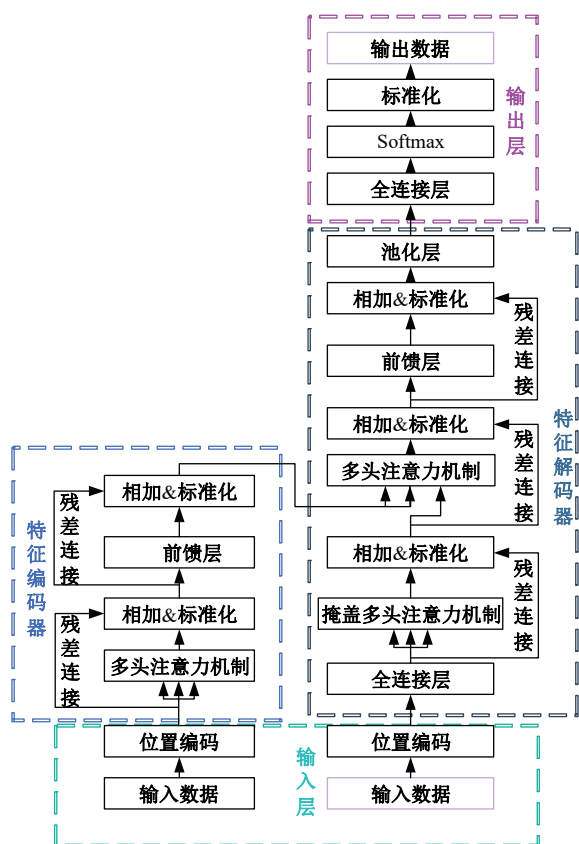


图 2 Transformer 结构图

Fig. 2 Transformer structure diagram

(1) 多头注意力机制

多头注意力机制是编码器的核心,其优点是可以减少对外部信息的依赖,通过计算特征间的相关性权重使重要的信息得到更好的表达。

多头自注意力机制是由多个自注意力机制通过拼接和线性变化得来。其计算过程如下:

$$\begin{cases} Q = W^q \cdot I \\ K = W^k \cdot I \\ V = W^v \cdot I \end{cases} \quad (11)$$

$$\text{Attention}(Q, K, V) = \text{Soft max} \left(\frac{K^T Q}{\sqrt{d}} \right) V \quad (12)$$

$$\text{MultiAttention}(Q, K, V) = \text{Concat}(H_1, H_2, \dots, H_n) \cdot W \quad (13)$$

$$H_i = \text{Attention}(IW_i^Q, IW_i^K, IW_i^V) \quad (14)$$

式中: Q 、 K 、 V 分别为查询矩阵、键矩阵和值矩阵; I 为特征矩阵; W^q 、 W^k 、 W^v 分别为 Q 、 K 、 V 的权重矩阵; d 为 K 的维度; Soft max 为归一化函数; Concat 函数为拼接多个自注意力头的输出; H_i 为第 i 个自注意力头; n 为多头注意力头数; W_i^Q 、 W_i^K 、 W_i^V 分别为第 i 个注意力头 Q 、 K 、 V 的权重矩

阵; W 为多头自注意力权重矩阵。

(2) 前馈网络

前馈网络位于每个子层中,由两层全连接层和一个非线性激活函数组成,通过对各位置的特征进行独立的非线性变换,增强模型的特征表达能力。

(3) 残差连接与层归一化

残差连接位于特征编码器和解码器的每一子模块之间,可以更好地优化权重矩阵。归一化处理可以加快训练速度,减少过拟合。

1.3.3 解码器

解码器由掩盖多头自注意力机制、编码-解码器多头注意力层、前馈网络和残差归一化层组成,其作用是将中间向量解码为输出标签序列。

(1) 掩盖多头自注意力机制

掩盖多头自注意力机制为解码器的核心部分,其在多头注意力机制的基础上加入了遮挡操作,使解码器在生成目标序列时,每个时间步只能访问已生成部分,避免模型获取未来信息。

(2) 编码-解码器多头注意力层

编码-解码器多头注意力层用于连接编码器和解码器之间的信息流,能够有效捕捉输入序列(编码器输出)与生成序列(解码器中间状态)之间的复杂依赖关系。

1.3.4 输出层

输出层为 Transformer 模型的最后部分,负责将模型处理后的高维特征映射到任务所需的具体结果形式。其包括全连接层、Softmax 层和正则化层。

1.4 基于 Hoyer 稀疏度改进的余弦相似度

1.4.1 Hoyer 稀疏度值

Hoyer 值是一种衡量向量稀疏性的指标,能够有效表征信号的能量分布特征,其在突出信号局部特征和异常点方面具有优势。计算过程如下:

$$H_x = \left(\sqrt{t} - \frac{\sum |x|}{\sqrt{\sum x^2}} \right) / (\sqrt{t} - 1) \quad (15)$$

式中: H_x 为 x 的 Hoyer 值; t 为 x 的长度; x 为振动信号。

综上, Hoyer 值通过结合 L1 和 L2 范数来度量向量的稀疏性,数值越接近 1, 向量越稀疏; 越接近 0, 向量越密集。

1.4.2 高斯径向基函数(GRBF)

高斯径向基函数(Gaussian radial basis function, GRBF)是一种基于高斯分布的径向基函数,用于衡量两个数据点之间的相似性。计算过程如下:

$$\text{GRBF}(x, y, \sigma) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (16)$$

式中: x 和 y 为原始样本; σ 为径向范围参数。

GRBF的输出取决于两个输入样本之间的欧几里得距离,距离越小,值越接近1;反之,则接近0。

1.4.3 余弦相似度

余弦相似度(Cosine Similarity)是一种常用于文本数据分析的相似度度量方法。它通过计算两个向量在空间中的夹角,来衡量两个向量之间的相似程度。计算过程如下:

$$\cos(\theta) = \frac{x \cdot y}{\|x\| \cdot \|y\|} \quad (17)$$

式中: θ 为余弦角。

余弦相似度值的范围为 $[-1, 1]$,两个向量相似度越高,余弦相似度越接近1;越不相似,值越接近-1;当值为0时,表示两个向量正交,即完全不相干。

1.4.4 Hoyer稀疏度改进的余弦相似度

如图3为基于Hoyer稀疏度改进的余弦相似

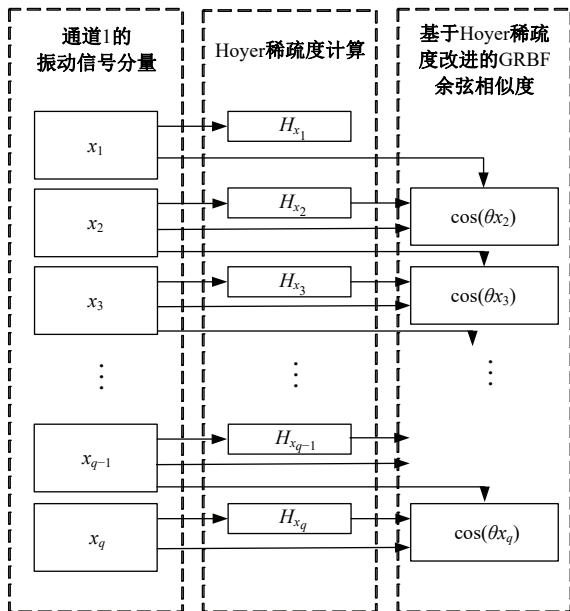


图3 基于Hoyer稀疏度改进的余弦相似度计算结构图

Fig. 3 Cosine similarity calculation structure based on improved Hoyer sparsity

度计算结构图。

通过高斯径向基函数(GRBF)和余弦相似度方法可以评价当前样本和历史样本之间的稳定性,采用Hoyer稀疏度值来提高GRBF的余弦相似度。Hoyer稀疏度改进的余弦相似度计算过程如下:

$$\cos(\theta_x) = \frac{\text{GRBF}(x, x_h, H_x)}{\sqrt{\text{GRBF}(x, x, H_x)} \sqrt{\text{GRBF}(x_h, x_h, H_x)}} \quad (18)$$

式中: θ_x 为 x 的改进余弦角; x_h 为 x 的历史样本。

2 MFTDCN模型介绍

2.1 模型结构

MFTDCN模型由数据级融合、特征提取和分类器组成。图4为MFTDCN模型的结构图。

2.1.1 数据级融合

首先,为了减少数据的维度和冗余,采用主成分分析(PCA)对8个通道的原始振动信号进行降维处理,得到3个主成分分量。

接下来,通过Hoyer稀疏度改进的余弦相似度方法来计算每个通道数据在融合过程中的权重。权重计算过程如下:

$$\omega_i = \frac{\cos(\theta_{x_i})}{\sum_{i=1}^n \cos(\theta_{x_i})} \quad (19)$$

式中: ω_i 为第 i 个样本的权值; n 为参与融合样本的个数。

由于第一个样本缺乏历史样本,所以第一个样本仅用Hoyer值加权。

在获得权重之后,将降维后的3个分量时域信号通过短时傅里叶变换(STFT)转换为时频图像。这一转换能够有效捕捉信号在不同时间和频率域的变化特征,为后续的故障诊断提供丰富的时频信息。

最后,利用之前计算得到的权重,将这3个时频图像加权融合成1个RGB图像。每个时频图像的权重决定了其在最终RGB图像中的贡献,从而实现了有效的数据级融合。此数据级融合方法能够综合各通道的优势,提高故障诊断系统的准确性和鲁棒性。

2.1.2 特征提取和分类

首先,将融合后的RGB图像输入到可变形卷积网络(DCN)中进行特征提取。可变形卷积

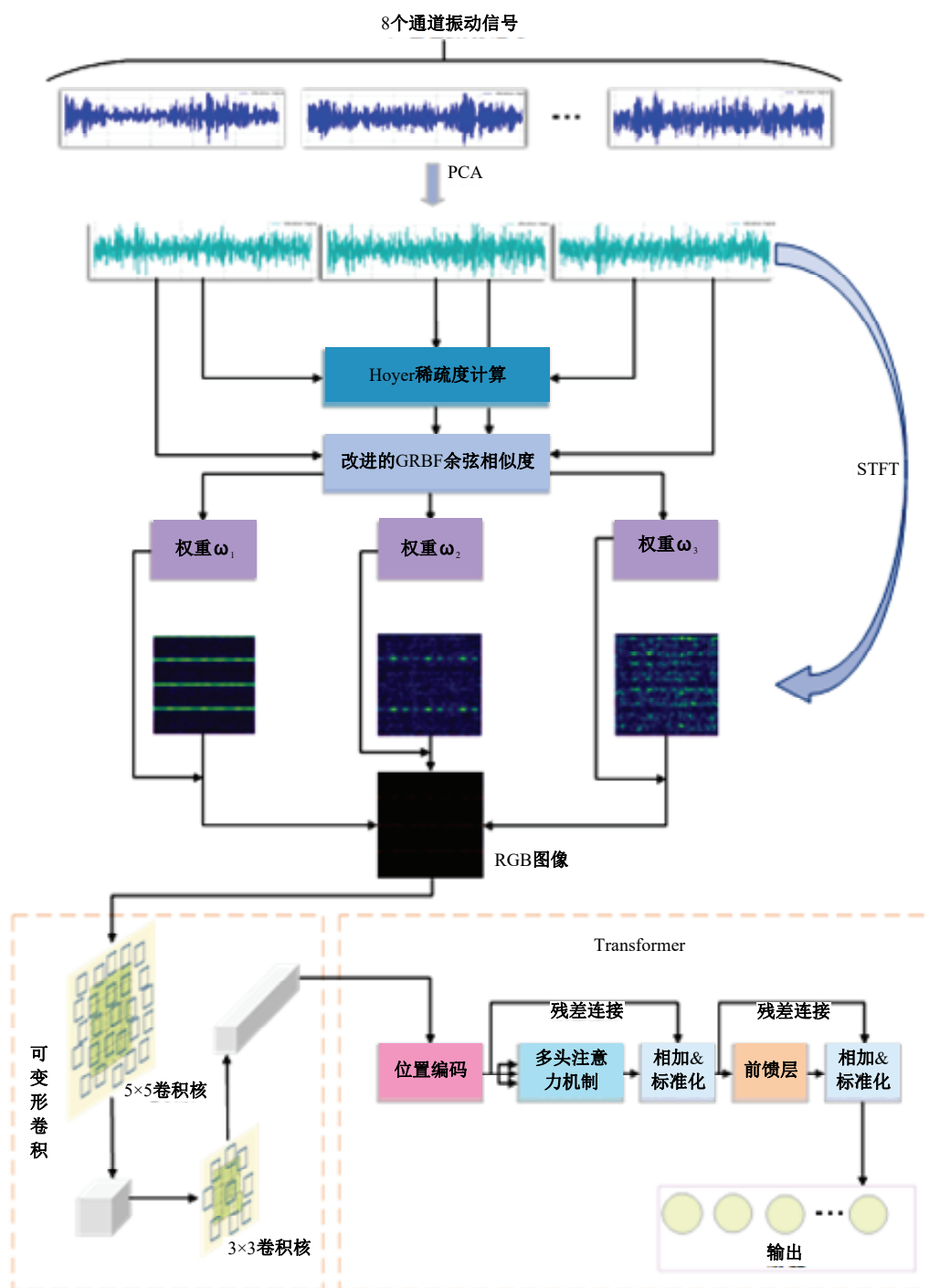


图4 MFTDCN模型图

Fig. 4 MFTDCN model diagram

能自适应地调整卷积核的采样位置,有效捕获图像中复杂的几何变形和局部特征,从而提高特征提取的灵活性和准确性。

接下来,将DCN提取到的初步特征传输至Transformer网络中进一步处理。Transformer网络可以使特征能够在全局范围内进行交互,从而提升特征的表达能力。

最后,将经过Transformer网络增强的特征

输入到分类器中进行分类,根据提取到的深度特征对输入数据进行准确的故障诊断。

此方法结合DCN和Transformer的特征提取方法,充分利用了可变形卷积在局部特征捕获方面的优势和Transformer在全局特征建模方面的能力,有效提高了模型在复杂场景下的分类性能和泛化能力。

2.2 故障诊断流程

该模型的训练过程包括以下步骤:首先进行数据级融合部分,然后进行训练集的特征提取,根据训练结果进行参数更新,随后将测试集提取的特征输入到更新后的模型中,最后在达到最大迭代次数时结束训练并输出结果。具体故障诊断流程如图 5 所示。

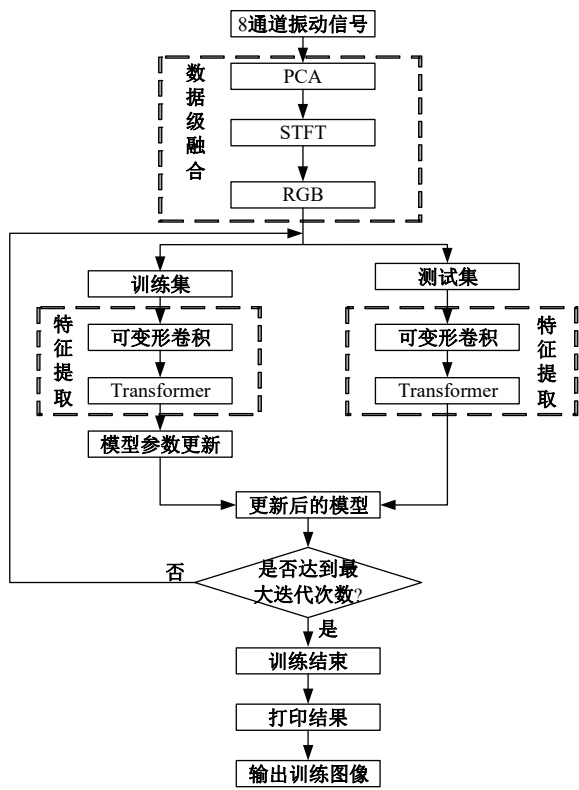


图 5 故障诊断流程
Fig. 5 Fault diagnosis process

3 实验验证与结果分析

在本节中,首先详细描述了用到的齿轮箱数据集;随后,为验证 MFTDCN 模型的正确性,进行了不同二维图像转换方法、不同权重计算方法的对比;接着,将 MFTDCN 模型与其他诊断方法进行整体对比;最后,又将训练数据与测试数据划分成不同比例进行实验,以此来检验模型性能。

3.1 实验数据

本实验首先采用中国东南大学提供的齿轮箱数据集进行模型验证,随后使用北京工业大学与北京交通大学联合公开的 WT 数据集^[19]对模型的泛化性能进行测试。

中国东南大学的齿轮箱数据集是从传动系动态模拟器(DDS)上收集的,包含轴承数据集和齿

轮数据集,每个子数据集又包含两种工况:转速 20 Hz(1 200 r/min)-负载 0V(0 N·m)和转速 30 Hz(1 800 r/min)-负载 2V(7.32 N·m)。所以可分为 bearingset_20_0、bearingset_30_2、gearset_20_0、gearset_30_2 四组数据集进行实验。轴承数据集有 4 种故障类型,共 5 种状态:滚动体故障、内圈故障、外圈故障、复合故障和健康状态。齿轮数据集也有 4 种故障类型,共 5 种状态:齿轮缺口、齿轮断齿、齿根裂纹、齿面磨损和健康状态。

每种故障类型有 8 个通道信号:电机振动信号、行星齿轮箱 3 个方向(X、Y、Z)振动信号、电机扭矩和平行齿轮箱 3 个方向(X、Y、Z)振动信号。将这 8 个通道的数据降维到 3 个分量后进行数据级融合。每种工况取 300 个样本,并按照 7:3 划分为训练集和测试集。具体数据信息如表 1 所示。

表 1 东南大学数据集中轴承和齿轮数据集划分表
Table 1 Division table of bearing and gearbox data sets in southeast university data set

轴承/齿轮箱	类型	训练集	测试集	标签
轴承数据	滚动体故障	210	90	0
	内圈故障	210	90	1
	外圈故障	210	90	2
	复合故障	210	90	3
	健康状态	210	90	4
齿轮数据	齿轮断齿	210	90	0
	齿轮缺口	210	90	1
	齿面磨损	210	90	2
	齿根裂纹	210	90	3
	健康状态	210	90	4

北京工业大学与北京交通大学联合公开的 WT 数据集包括 20、25、30、35、40、45、50、55 Hz,共 8 种不同转速工况,每种工况包含齿轮健康、齿轮损坏、齿面磨损、齿根断裂和缺齿共 5 类状态。每组数据均提供 X 轴和 Y 轴两个方向的振动信号,采样频率为 48 kHz。

模型在每个任务上进行了 100 epoch 的训练,采用 Adam 优化器,学习率为 0.001,权重衰减为 0.000 1。为了使实验更具有说服力和公平性,每种方法在每个数据集上都进行 3 次测试,并取每次测试中最后 10 次测试结果的平均值作为测试最终结果。

3.2 数据级融合部分方法对比

此方法的核心在于数据融合部分,其中选择合适的图像变换方法和权重计算方法至关重要。本文首先将所采用的短时傅里叶变换(STFT)与

其他常见的图像变换方法进行对比,验证了ST-FT 在该场景下的优越性。接着,比较了利用Hoyer稀疏度改进的余弦相似度计算权重的方法与其他权重计算的方法,证明了本文使用的权重计算方法在数据融合过程中计算通道分配权重的合理性。

3.2.1 图像变换方法对比

将短时傅里叶变换(STFT)与灰度图(Gray)、格拉姆角场(GAF)、马尔科夫(MTF)变换方法进行对比,得到的图像如图6所示。为了更清楚地看到图像的变化,将8个通道的振动信

号分别转换为二维图像,与振动信号降维之后再转换为二维图像的三个图像对比,可以看出,振动信号经过PCA降维后再转为二维图像可以有效地保留信号的主成分,同时压缩了原始数据中的冗余信息。将3个图像经过加权融合后得到RGB图像,融合后的RGB图像也成功地保留了数据中的重要成分。经过仔细观察,STFT生成的RGB图像在视觉上具有更高的辨识度。相较于其他变换方法,STFT在时频图像中能够展现出更明显的信息,更清晰地区分信号中的不同频率成分,因此更适合用于信号分析和故障诊断。

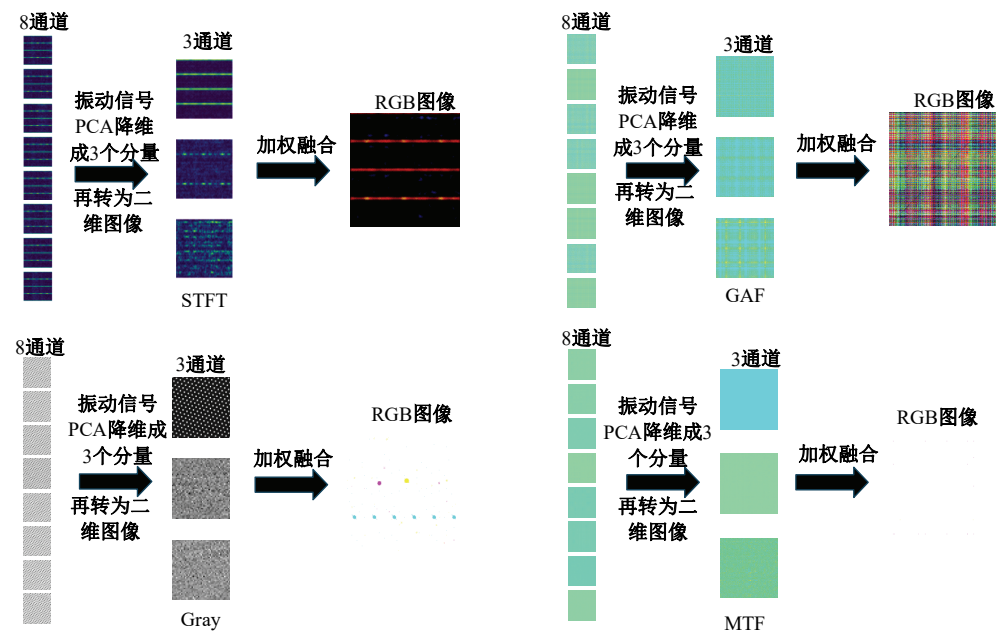


图 6 不同图像转换方法的对比效果图

Fig. 6 Comparative effect of different image conversion methods

在本文提出的MFTDCN诊断模型中,只改变图像转换方法,保证其他所用方法不变的情况下,故障诊断的准确率如表2所示。可以看出,STFT方法的效果最为优越,平均准确率达到99.46%,明显高于其他方法。这一结果与通过视觉观察得到的结论一致,进一步验证了STFT在该任务中的显著优势。此结果表明,STFT能够更有效地捕捉信号中的关键信息,从而提升数据分析和故障诊断的精度。

如表3所示,展示了在bearingset_30_2和gearset_30_2工况下,不同故障类型的三个分量时频图与生成的RGB图像的对比。从中可以看出,每种故障类型的RGB图像具有明显的差异,这为后续的特征提取奠定了坚实的基础。

表 2 不同图像转换方法准确率对比

Table 2 Comparison of accuracy rates of different image conversion methods

工况	Gray	GAF	MTF	STFT
bearingset_20_0	82.16	96.16	74.24	100.00
bearingset_30_2	84.22	93.76	84.22	99.92
gearset_20_0	86.18	88.71	82.11	98.27
gearset_30_2	94.91	93.80	79.67	99.65
均值	86.87	93.11	80.06	99.46

3.2.2 计算权重方法对比

计算出合适的权重能确保融合后的RGB图像最大程度地保留关键信息,因此权重的计算也很关键。计算权重的方法对比依旧是在只改变计算权重方法,其他所用方法不变的情况下,将Hoyer稀疏度改进的余弦相似度方法与信息熵、方差权重、未改进的余弦相似度和能量权重四种

表 3 不同故障类型的 3 个分量时频图与生成的 RGB 图像对比

Table 3 Comparison of time-frequency graphs of three components of different fault types with generated RGB images

类型	工况 bearingset_30_2					工况 gearset_30_2				
	滚动体	内圈	外圈	复合	正常	缺口	断牙	裂纹	磨损	正常
分量 1										
分量 2										
分量 3										
RGB										

常用的权重计算方法进行了对比分析。不同权重计算方法的准确率对比结果如表 4 所示。

通过对比实验可以发现,在使用 STFT 进行变换时,不同的权重计算方法对准确率产生了不同程度的影响。尽管 5 种方法的准确率差异不大,但可以看出,Hoyer 稀疏度改进的余弦相似度方法相较于其他方法略微表现更优,展现出一定的优势。

通过上述两个对比试验,验证了本方法中采用的 STFT 变换以及 Hoyer 稀疏度改进的余弦相似度计算融合权重在性能上的优势。这些结果充分证明了该方法的有效性和正确性。

3.3 东南大学数据集模型对比实验结果分析

3.3.1 实验对比

为进一步验证模型在故障诊断方面的性能,本实验将所提的 MFTDCN 方法与 RGB-

ResNet^[20]、Gray-LeNet-5^[21]、Raw-TICNN^[22]、Matrix-CNN^[23]、CAM_MCFCNN^[24] 五种方法进行对比。结果如表 5 所示,其中,RGB-ResNet、Gray-LeNet-5、Raw-TICNN、Matrix-CNN 四种方法的准确率来自参考文献[25]。

对模型的复杂度进行分析,其中 RGB-ResNet、Gray-LeNet-5、Raw-TICNN、Matrix-CNN 和 CAM_MCFCNN 模型的参数量分别为 25 636 712^[25]、97 408^[25]、90 772^[25]、2 117 120^[25]和 118 143,而本文提出的 MFTDCN 模型参数量为 316 019,相较而言处于适中的水平。在 Intel Core i5 处理器、NVIDIA GeForce RTX 3050 显卡的计算机上,采用 PyCharm 2023. 2. 1 作为开发环境,MFTDCN 模型的完整运行时间约为 181 s。该时间包含模型训练与测试全过程,在保持较低计算资源消耗的同时,展现出较高的运行效率。

从表 5 中可以看出,MFTDCN 模型在故障诊断任务中的表现优于其他对比方法,平均准确率分别高出 2.46%、5.65%、0.34%、1.84% 和 2.57%。此外,该模型在所有工况下均表现出色,特别是在 bearingset_20_0 工况中准确率达到 100%。上述结果充分表明,MFTDCN 方法能够更有效地提取和利用复杂故障特征信息,从而显著提高诊断的准确性和稳定性。

为了观察 MFTDCN 模型的训练稳定性和收敛性,在 4 种不同工况下分别绘制了训练准确率

表 4 不同权重计算方法的准确率对比

Table 4 Comparison of accuracy of different weight calculation methods

工况	信息熵	方差权重	未改进的余弦相似度	能量权重	Hoyer 稀疏度改进的余弦相似度
bearingset_20_0	99.42	99.83	99.87	99.38	100.00
bearingset_30_2	99.17	99.63	100.00	99.55	99.92
gearset_20_0	94.02	93.94	96.24	95.08	98.28
gearset_30_2	98.89	98.94	99.48	98.86	99.65
均值	97.88	98.09	98.90	98.22	99.46

表 5 不同模型的故障诊断结果对比表

Table 5 Comparison table of fault diagnosis results of different models

工况	RGB-ResNet	Gray-LeNet-5	Raw-TICNN	Matrix-CNN	CAM_MCFCNN	MFTDCN
bearingset_20_0	99.84	95.12	99.84	99.55	95.11	100.00
bearingset_30_2	99.55	99.28	97.60	99.94	94.89	99.92
gearset_20_0	92.27	89.33	99.05	95.76	98.67	98.27
gearset_30_2	96.32	91.52	99.97	95.23	98.88	99.65
均值	97.00	93.81	99.12	97.62	96.89	99.46

曲线,如图7所示。由图7可以看出,经过20次迭代后,每种工况下的准确率均达到了最高值并保持平稳,并且每种工况下训练集的准确率都稳定在100%。尽管在 gearset_20_0 工况下,训练初期

测试集准确率较低且波动较大,但随着训练次数的增加,准确率逐渐稳定并趋于平稳。这表明,MFTDCN模型在不同工况下均能较快收敛,并维持较高且稳定的准确率。

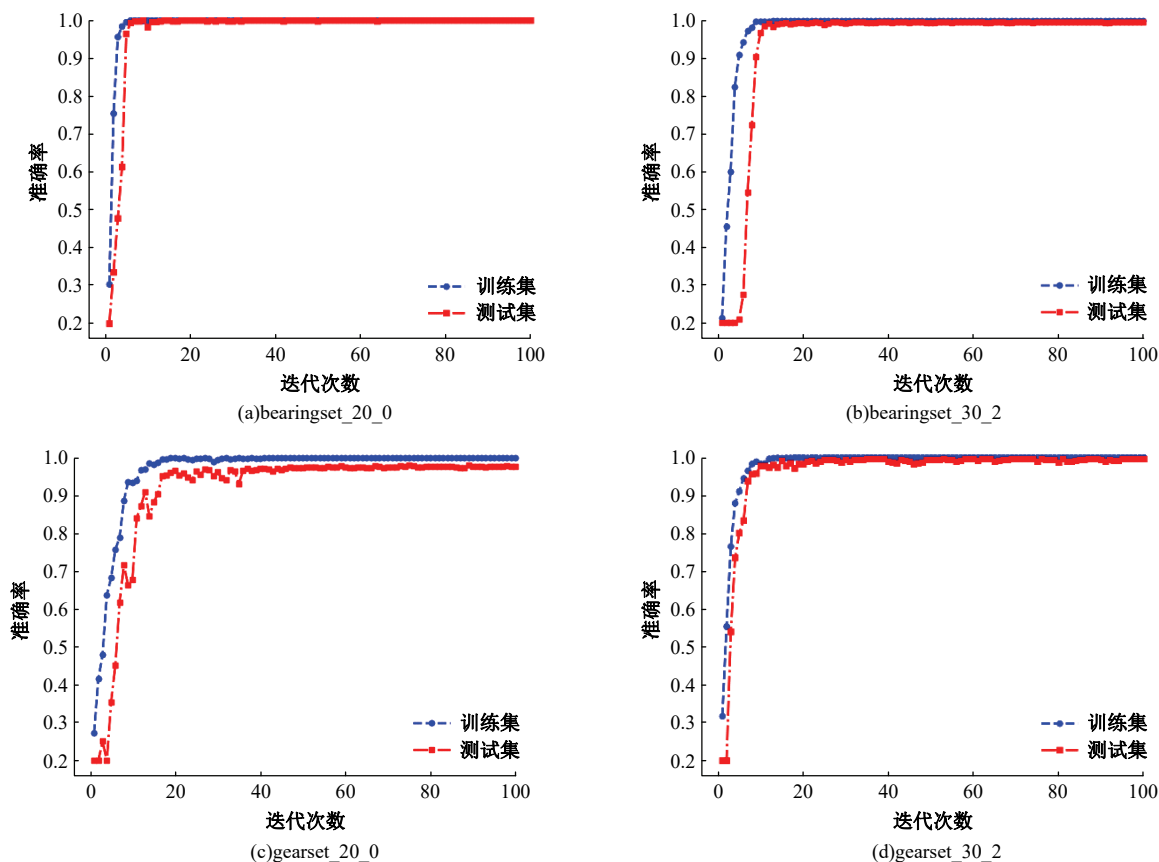


图7 MFTDCN方法各工况训练准确率曲线

Fig. 7 Training accuracy curve of MFTDCN method in each condition

图8展示了MFTDCN方法在4种工况下的混淆矩阵。从结果可以看出,在 gearset_20_0 工况下分类准确率略低,主要原因可能在于标签4与标签0的振动信号特征较为相似,导致模型在识别时容易将标签4误判为标签0。尽管存在一定的误分类情况,整体来看,MFTDCN方法在各类工况下依然表现出良好的分类效果,展现了较强的故障识别能力与泛化性能。

图9展示了 bearingset_20_0 工况下训练前后的测试集 t-SNE 特征分布情况,由两图对比可以看出,在没训练之前,测试集的5种类别的 t-SNE 特征分布非常杂乱没有规律,训练后的5种类别的 t-SNE 特征分布为明显的5个部分,证明 MFTDCN 可以准确地将5种类别的特征分布区分开。

3.3.2 小样本试验

为评估 MFTDCN 模型在训练样本不足时的表现,本实验将训练和测试样本的比例分别调整

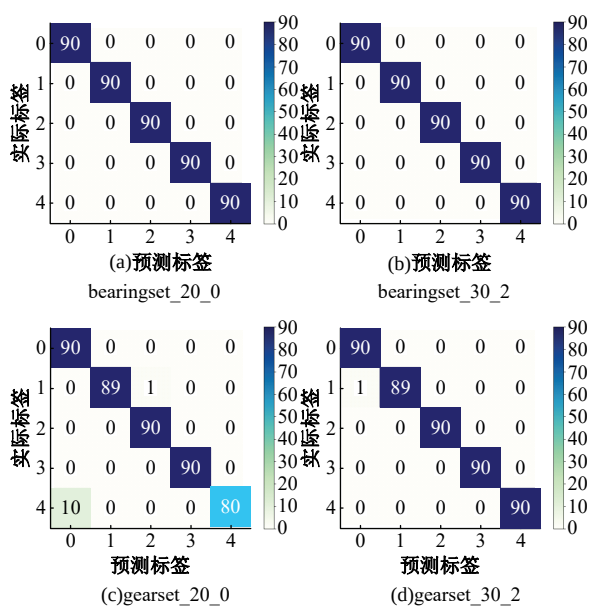


图8 4种工况下的混淆矩阵

Fig. 8 Confusion matrix under four working conditions

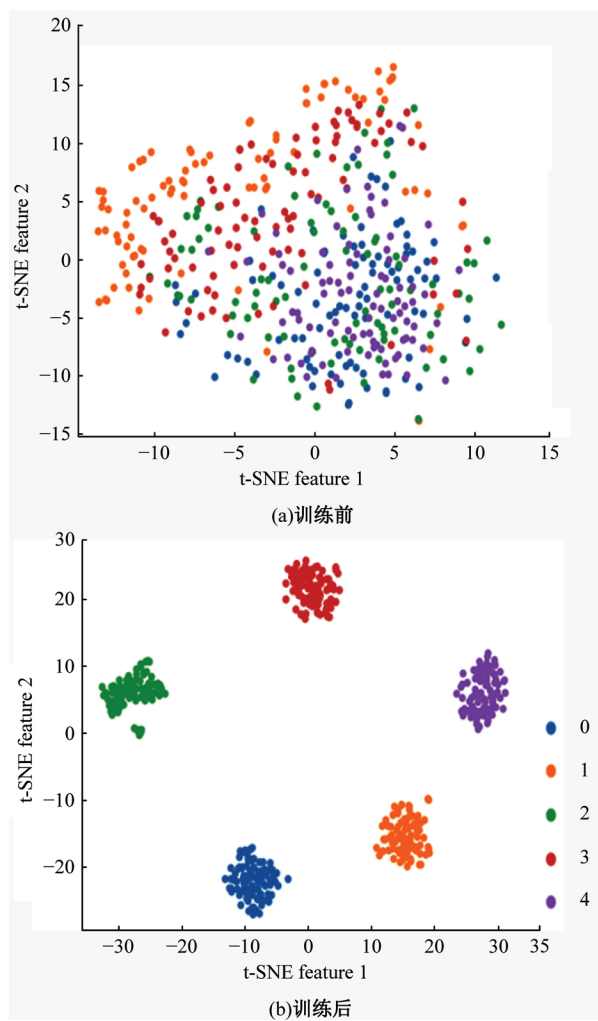


图 9 训练前后测试集 t-SNE 特征分布对比

Fig. 9 Comparison of t-SNE feature distribution in test set before and after training

为 1:9、2:8 和 3:7。实验结果如表 6 所示,尽管训练样本数量有限,模型的平均准确率仍然保持在 95% 以上,尤其是在 bearingset_20_0 和 bearingset_30_2 工况下表现尤为突出。这表明,MFTDCN 模型能够有效应对训练样本不足的情况,验证了其在小样本环境下的优越性能和应用潜力。

图 10 为训练和测试样本的比例为 2:8 情况下,MFTDCN 在各工况的训练曲线。可以发现,训练样本减少时,模型达到平稳状态所需的

表 6 小样本训练准确率

工况	1:9	2:8	3:7	7:3
bearingset_20_0	98.12	99.36	99.59	100.00
bearingset_30_2	97.16	99.27	99.47	99.92
gearset_20_0	90.25	92.96	93.55	98.27
gearset_30_2	94.48	96.58	97.74	99.65
均值	95.00	97.04	97.59	99.46

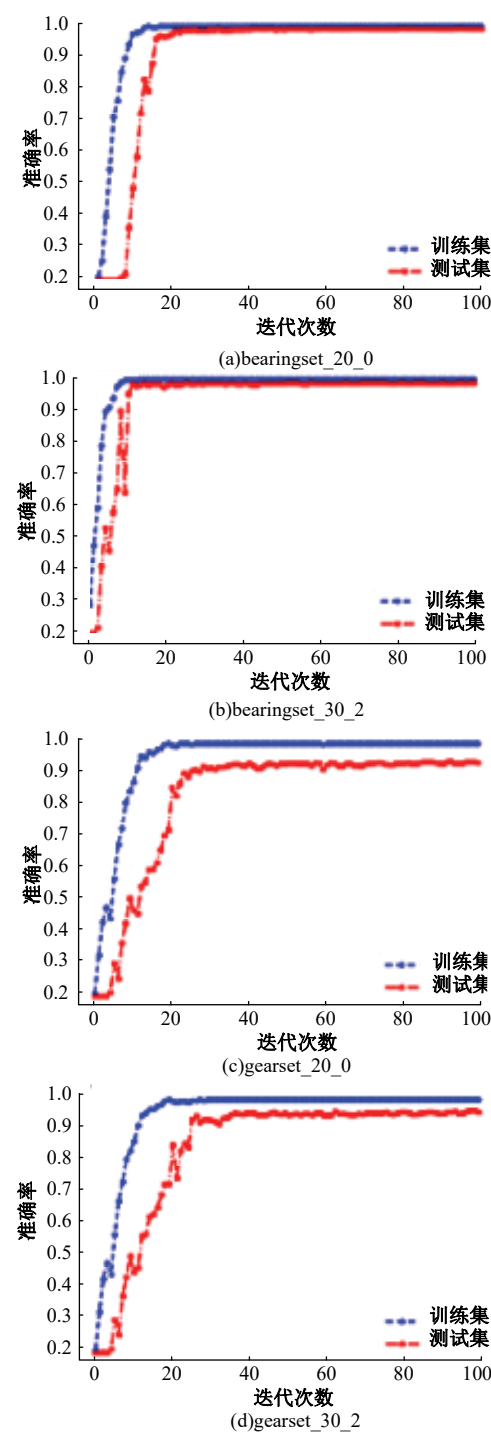


图 10 小样本实验各工况训练曲线

Fig. 10 Training curve of each condition in small sample experiment

训练次数增加。尽管在训练样本不足情况下的准确率较训练样本充足时有所降低,但其准确率依然保持在较高水平,表现出良好的收敛性。

3.3.3 轴承和齿轮数据集合并实验分析

为验证本方法能否根据振动信号准确识别轴承或齿轮的具体故障部位,现将轴承和齿轮在相

同工况下采集的数据合并为一个完整数据集,并统一设置标签为 0~8,其中标签“0”代表正常状态的数据,因此该类别的样本量为其他类别的两倍,通过数据合并操作,数据集就只有 20_0 和 30_2 两种工况,此实验能实现对整个齿轮箱故障位置的精确识别。

合并数据集的识别准确率如表 7 所示,分别达到了 98.26% 和 98.38%,表现出较高的识别精度,表明 MFTDCN 具备良好的齿轮箱故障诊断能力。

从图 11 的训练准确率曲线可知,两个数据集的准确率在初期均出现了较大的波动,但随着训练次数的增加逐渐趋于稳定。其中,20_0 工况准确率波动较大的主要原因可能是该工况下的数据集 gear-set_20_0 更为复杂,数据特征更难区分,从而导致模型在训练初期的分类准确率波动较明显。

图 12 为合并数据集后 30_2 工况下的混淆矩阵,可以看出,MFTDCN 模型在合并数据集上的分

表 7 合并数据集准确率表

Table 7 Accuracy table of merged datasets	
工况	准确率/%
20_0	98.26
30_2	98.38

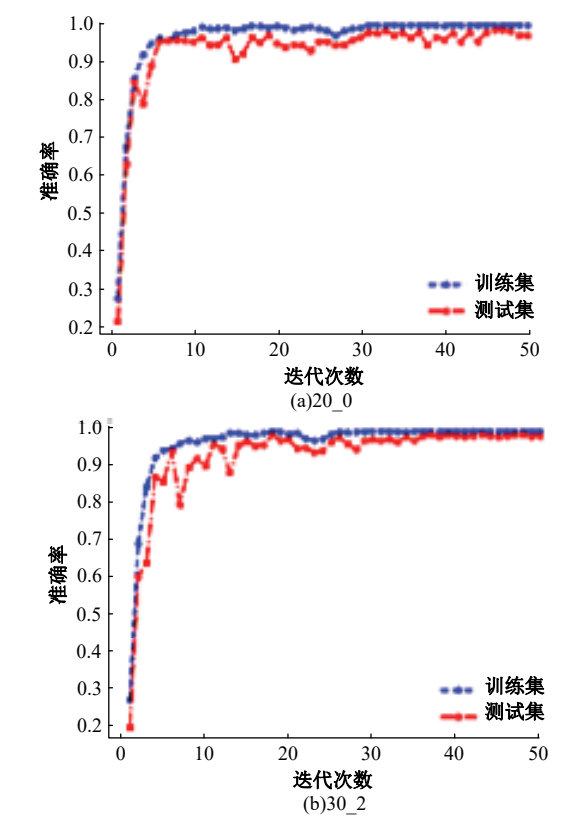


图 11 合并数据集训练准确率曲线
Fig. 11 Training accuracy curve of the merged dataset

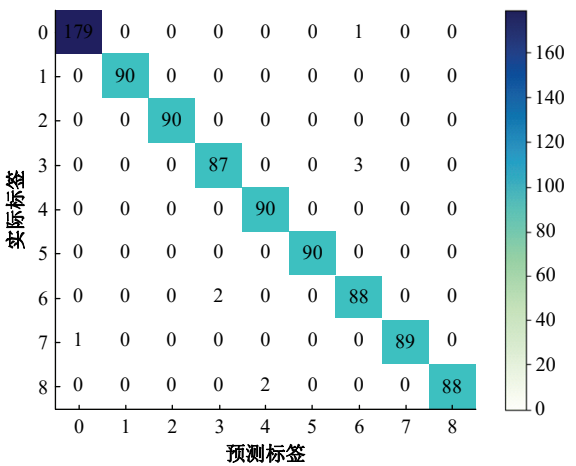


图 12 合并数据集 30_2 工况下混淆矩阵

Fig. 12 Confusion matrix under condition 30_2 with merged dataset

类表现较为理想。这种数据合并方式不仅提升了故障诊断任务的实际意义,同时也有效地验证了模型对具体故障部位和故障类型的准确识别能力。

3.4 泛化能力实验分析

在北京工业大学和北京交通大学公开数据集上经过 100 epoch 的训练,实验结果如表 8 所示。平均准确率为 98.8%,说明模型具有较强的泛化能力。

图 13 为 1_55 工况下的训练准确率曲线,可以

表 8 泛化能力实验准确率表

Table 8 Accuracy rate table for generalization ability experiment			
工况	准确率/%	工况	准确率/%
1_20	99.3	1_40	98.5
1_25	99.6	1_45	98.3
1_30	98.6	1_50	98.1
1_35	98.2	1_55	99.6

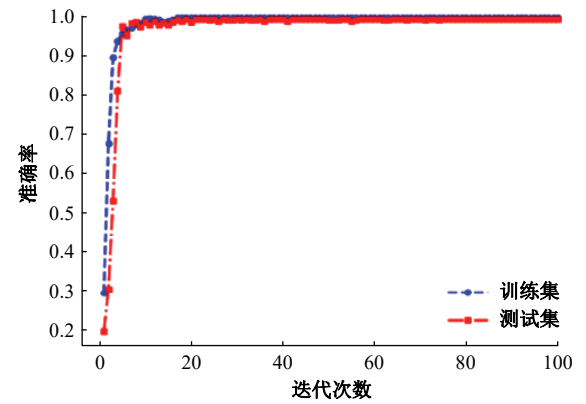


图 13 泛化能力实验 1_55 工况下训练曲线
Fig. 13 Training curve under condition 1_55 in generalization experiment

看出,模型准确率整体表现非常平稳,几乎没有明显波动,充分体现了MFTDCN模型在不同数据集上的良好适应性与稳定性,说明其具备较强的泛化能力。

4 结 论

(1)在数据预处理部分,采用PCA降维将8个通道的振动信号降维为3个分量,采用短时傅里叶变换,将一维振动信号转换为二维时频图像。该方法在保留主要成分的同时,去除了冗余信息,突出并增强了关键成分。

(2)在数据级融合部分,用Hoyer稀疏度改进的余弦相似度计算3个分量的权重,并利用权重将3个分量的时频图融合为1个RGB图像。此过程进一步降低了数据复杂度,同时增强了重要信息的表达。

(3)通过对比实验验证了该方法在数据级融合方面的优势。在小样本情况下,模型仍能保持较高的准确率和稳定性,证明其在训练样本不足时的有效性和适应性。

参考文献:

- [1] Saufi S R, Ahmad Z A B, Leong M S, et al. Gearbox fault diagnosis using a deep learning model with limited data sample[J]. IEEE Transactions on Industrial Informatics, 2020, 16(10): 6263-6271.
- [2] 侯召国,王华伟,熊明兰,等. 基于迁移学习与加权多通道融合的齿轮箱故障诊断[J]. 振动与冲击, 2023, 42(9): 236-246.
Hou Zhao-guo, Wang Hua-wei, Xiong Ming-lan, et al. Gearbox fault diagnosis based on transfer learning and weighted multi-channel fusion[J]. Journal of Vibration and Shock, 2023, 42(9): 236-246.
- [3] She D M, Yang Z C, Duan Y D, et al. A meta transfer learning-driven few-shot fault diagnosis method for combine harvester gearboxes[J]. Computers and Electronics in Agriculture, 2024, 227: 109605.
- [4] 王进花,刘秦玮,曹洁,等. 基于SCACGAN的小样本齿轮箱故障诊断[J]. 北京航空航天大学学报, 2026, 52(3): 713-723.
Wang Jin-hua, Liu Qin-wei, CAO Jie, et al. Fault diagnosis of gearbox with small-sample based on SCACGAN[J]. Journal of Beijing University of Aeronautics and Astronautics, 2026, 52(3): 713-723.
- [5] 刘辉,李阳,侯一民. 用于轴承故障诊断任务的轻量化卷积网络[J]. 哈尔滨理工大学学报, 2024(4): 80-88.
- [6] Liu Hui, Li Yang, Hou Yi-ming. Lightweight convolutional network for bearing fault diagnosis[J]. Journal of Harbin University of Science and Technology, 2024(4): 80-88.
- [7] Lei Y G, Jia F, Lin J, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 3137-3147.
- [8] Hu Q, Si X S, Zhang Q H, et al. A rotating machinery fault diagnosis method based on multi-scale dimensionless indicators and random forests[J]. Mechanical Systems and Signal Processing, 2020, 139: 106609.
- [9] 石永芳,徐庆宏,姜宏,等. 基于特征差异性学习卷积神经网络的齿轮箱故障诊断方法[J]. 机床与液压, 2023, 51(24): 176-183.
Shi Yong-fang, XU Qing-hong, JIANG Hong, et al. Gearbox fault diagnosis method based on feature difference learning convolutional neural network[J]. Machine Tool & Hydraulics, 2023, 51(24): 176-183.
- [10] Zhao D Z, Cui L L, Liu D D. Bearing weak fault feature extraction under time-varying speed conditions based on frequency matching demodulation transform[J]. IEEE/ASME Transactions on Mechatronics, 2022, 28(3): 1627-1637.
- [11] Goyal D, Dhama S S, Pabla B S. Non-contact fault diagnosis of bearings in machine learning environment[J]. IEEE Sensors Journal, 2020, 20(9): 4816-4823.
- [12] Zhang N N, Wu L F, Yang J, et al. Naive bayes bearing fault diagnosis based on enhanced independence of data[J]. Sensors, 2018, 18(2): 18020463.
- [13] Zhao X L, Jia M P, Liu Z. Semisupervised deep sparse auto-encoder with local and nonlocal information for intelligent fault diagnosis of rotating machinery[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 1-13.
- [14] Wang X H, Meng R X, Wang G T, et al. The research on fault diagnosis of rolling bearing based on current signal CNN-SVM[J]. Measurement Science and Technology, 2023, 34(12): 125021.
- [15] Rezaeianjouybari B, Shang Y. Deep learning for prognostics and health management: state of the art, challenges, and opportunities[J]. Measurement, 2020, 163: 107929.
- [16] Zhu X X, Hou D N, Zhou P, et al. Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images[J]. Measurement, 2019,

- 138: 526-535.
- [16] Wang S Y, Tian J Y, Liang P F, et al. Single and simultaneous fault diagnosis of gearbox via wavelet transform and improved deep residual network under imbalanced data[J]. *Engineering Applications of Artificial Intelligence*, 2024, 133: 108146.
- [17] Li C Y, Hu Y H, Jiang J W, et al. Fault diagnosis of a marine power-generation diesel engine based on the Gramian angular field and a convolutional neural network[J]. *Journal of Zhejiang University-SCIENCE A*, 2024, 25(6): 470-482.
- [18] Ding X X, He Q B. Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis[J]. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(8): 1926-1935.
- [19] Liu D D, Cui L L, Cheng W D. A review on deep learning in planetary gearbox health state recognition: methods, applications, and dataset publication[J]. *Measurement Science and Technology*, 2023, 35(1): 012002.
- [20] Wen L L, Li X Y, Gao L. A transfer convolutional neural network for fault diagnosis based on ResNet-50[J]. *Neural Computing and Applications*, 2020, 32(10): 6111-6124.
- [21] 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*, 2017, 65(7): 5990-5998.
- [22] Zhang W, Li C H, Peng G L, et al. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load[J]. *Mechanical Systems and Signal Processing*, 2018, 100: 439-453.
- [23] Xia M, Li T, Xu L, et al. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks[J]. *IEEE/ASME Transactions on Mechatronics*, 2017, 23(1): 101-110.
- [24] Li H M, Huang J Y, Gao M J, et al. Multi-view information fusion fault diagnosis method based on attention mechanism and convolutional neural network[J]. *Applied Sciences*, 2022, 12(22): 11410.
- [25] Xie T L, Huang X F, Choi S K. Intelligent mechanical fault diagnosis using multisensor fusion and convolution neural network[J]. *IEEE Transactions on Industrial Informatics*, 2021, 18(5): 3213-3223.