

DOI: 10.13382/j.jemi.B2407691

应用于齿轮箱故障诊断的小样本图像生成方法*

高文超¹ 陈一帆¹ 陈诗雨¹ 周思杰¹ 黄俊²

(1. 中国矿业大学(北京)人工智能学院 北京 100083;

2. 安徽省工业互联网智能应用与安全工程研究中心 马鞍山 243023)

摘要:行星齿轮箱是一种广泛应用于工业领域的关键传动装置,其在复杂工况和长期负荷下易出现故障。传统的故障诊断方法依赖于专家经验和昂贵设备,存在数据稀缺和诊断效率低的问题。针对这一挑战,近年来生成对抗网络(GAN)的发展为图像生成和数据增强提供了新的解决方案。然而,现有GAN模型在处理小样本数据时,常出现语义错位和伪影问题,限制了其在智能故障诊断领域的应用潜力。为此,提出了一种基于多尺度渐进式特征融合的生成对抗网络(MSA-PF-GAN)模型,通过引入渐进式解码器结构与多尺度注意力模块,有效提升小样本条件下的图像生成质量及故障诊断精度。实验基于两个独立的行星齿轮箱故障数据集进行验证,结果显示,该方法显著降低了生成图像的FID分数,提升了诊断准确率(分别提高35%和20%)。在多种评价指标上,MSA-PF-GAN均优于其他主流方法。进一步分析表明,该模型通过渐进式特征融合和多尺度注意机制,不仅在生成图像的多样性和真实感上表现优异,还能有效增强对复杂故障特征的捕捉能力。因此,该技术在行星齿轮箱故障诊断领域具有有效的应用潜力和实际价值。

关键词: 图像生成;生成对抗网络;数据增强;齿轮箱故障诊断

中图分类号: TH132;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 510.70

Small sample image generation method applied to gearbox fault diagnosis

Gao Wenchao¹ Chen Yifan¹ Chen Shiyu¹ Zhou Sijie¹ Huang Jun²

(1. School of Artificial Intelligence, China University of Mining and Technology (Beijing), Beijing 100083, China;

2. Anhui Industrial Internet Intelligence Application and Security Engineering Research

Center, Ma'anshan 243023, China)

Abstract: Planetary gearboxes are widely used as essential transmission devices in industrial applications, yet they are prone to failures under complex operating conditions and prolonged loads. Traditional fault diagnosis methods heavily rely on expert knowledge and expensive equipment, facing challenges such as data scarcity and low diagnostic efficiency. To address these limitations, the development of generative adversarial networks (GANs) has provided innovative solutions for image generation and data augmentation in recent years. However, existing GAN models often encounter issues such as semantic misalignment and artifacts when processing small-sample datasets, limiting their potential in intelligent fault diagnosis. In this context, this paper proposes a multi-scale attention and progressive feature fusion GAN (MSA-PF-GAN) model, which integrates a progressive decoder structure with multi-scale attention mechanisms to significantly improve image generation quality and fault diagnosis accuracy under small-sample conditions. Experiments conducted on two independent planetary gearbox fault datasets validate the proposed method, showing that it substantially reduces the FID score and enhances diagnostic accuracy (by 35% and 20%, respectively). Across multiple evaluation metrics, the MSA-PF-GAN outperforms other state-of-the-art methods. Further analysis demonstrates that the model, through progressive feature fusion and multi-scale attention mechanisms, excels in generating diverse and realistic images while effectively capturing complex fault features.

收稿日期: 2024-07-16 Received Date: 2024-07-16

* 基金项目: 中央高校基本科研业务费(2024ZKPYZN01)、安徽省工业互联网智能应用与安全工程研究中心开放基金(IASII24-09)、北京市高等教育学会(MS2022314)、煤炭行业高等教育国家研究项目(2021MXJG44)、教育部产学研合作协同育人项目(202102210008)资助

Therefore, this technique shows promising potential and practical value in the field of planetary gearbox fault diagnosis.

Keywords: image generation; generative adversarial network; data enhancement; gearbox fault diagnosis

0 引言

行星齿轮箱以其高效性和紧凑性广泛应用于现代工业系统的多个关键领域,如风力发电、矿山机械和航空航天等。然而,这些系统对高性能的要求及其复杂的工作环境增加了发生故障的风险,从而对设备的可靠性和安全性构成了严重威胁。因此,准确而迅速地对行星齿轮箱进行故障诊断,对于确保这些关键行业的平稳运行至关重要^[1]。

齿轮箱故障诊断的任务通常涉及识别和分类多种故障类型,包括齿轮磨损、裂纹和断齿等。传统的故障诊断方法通常依赖振动分析^[2]、声发射检测^[3]和油液分析^[4]等物理传感器和数据处理技术,这些方法虽然能够提供一定程度的故障识别能力,但在获取充足的故障样本方面存在挑战。特别是在数据稀缺的情况下,传统方法难以有效应用,因为它们需要大量的标注数据和复杂的数据处理流程。此外,这些方法在受到现场环境影响时,往往表现出不稳定性。

近年来,随着人工智能技术的发展,尤其是深度学习在故障诊断中的应用,为解决这一问题提供了新的思路。深度学习方法通过自动提取和学习大规模数据中的特征,可以显著提高故障诊断的准确性和效率^[5]。例如,卷积神经网络(convolutional neural network, CNN)已被广泛应用于故障图像的特征提取和分类,而长短期记忆网络则用于处理时间序列数据,从而捕捉齿轮箱运行过程中细微的异常变化^[6],多尺度注意力机制通过在不同尺度上对特征进行加权处理,能够进一步提升模型对复杂故障模式的识别能力^[7]。因此,将振动数据转化为时频图,使得故障特征能够以图像形式更直观地呈现,从而能够充分利用CNN等深度学习模型在图像识别中的强大特征提取能力。通过这种图像化的处理方式,模型可以自动识别并分类不同的故障类型,而无需依赖专家经验的特征提取。这不仅提高了诊断的自动化水平,还在故障特征提取的精度和泛化能力上取得显著提升。

尽管深度学习技术在故障诊断中取得了显著进展,但现有生成对抗网络(generative adversarial network, GAN)模型在小样本数据集上常面临生成图像质量和多样性不佳的挑战,传统方法难以有效地进行数据增强和模型训练^[8]。小样本数据集的数据量有限、难以通过传统方法获取足够训练样本的特殊数据集。在此类数据集中,模型面临较大的过拟合风险,同时难以提取出具有普遍性的特征。因此,研究如何在小样本数据集上实现高

质量的图像生成对于增强故障诊断模型的鲁棒性具有重要意义。为此,研究人员提出了多种创新的GAN变体,以生成高质量的故障图像,从而提升故障诊断的性能^[9]。

例如,FIGR^[10]尝试将元学习技术与GAN结合,以掌握来自同一类别的少量图像数据的分布,但其生成的图像在质量和多样性方面仍有待提升。DAGAN^[11]基于单一条件图像生成多样化图像,但在融合同类多图像信息方面表现受限。MatchingGAN^[12]通过结合对抗性学习和匹配过程生成更多样化且逼真的图像,尽管如此,该方法在处理复杂自然图像的融合上仍面临挑战。F2GAN^[13]通过融合条件图像的高级特征以增强图像细节填充能力,虽然在融合能力上取得了进展,但在生成复杂自然图像方面仍需进一步优化。此外,LoFGAN^[14]作为一种旨在解决F2GAN在语义对齐和多样性方面问题的局部融合生成对抗网络,虽然在一定程度上改善了上述问题,但在融合过程中会产生锯齿状伪影且局部位置固定性也会导致多样性损失。ILoFGAN^[15]强调极少量样本的高效利用和局部特征的融合,采用嵌入多头注意力(multi-head attention, MHA)模块的生成器来提高特征融合过程的准确性和灵活性,将时域波形转换为时频图,以凸显故障特征。使用局部融合模块充分利用极少量样本并融合局部特征。最后构建了一个嵌入多头注意力模块的新生成器,以有效提高特征融合过程的准确性和灵活性。

针对这一挑战,本文提出了一种基于多尺度和渐进式特征融合的生成对抗网络模型(MSA-PF-GAN)。该模型通过渐进式解码特征融合策略和多尺度特征融合技术,旨在提高少样本故障诊断的精度和可靠性。具体而言,MSA-PF-GAN通过在不同层次上将特征经过上采样处理后逐级注入网络结构,并与多尺度特征融合,从而生成高质量的故障图像,增强图像的多样性和真实感。

通过在两个独立的行星齿轮箱故障数据集上进行实验验证,MSA-PF-GAN在故障诊断任务中显示了优异的性能,显著提高了故障诊断的准确率,并且在图像清晰度和多样性方面表现突出。这表明该模型在实际工业应用中具有广泛的潜力和应用价值。

1 理论基础

GAN近年来被广泛应用于数据增强和稀有样本生成领域,在齿轮箱故障诊断中,通过GAN生成多样化的故障样本,可以有效提升诊断的准确性和鲁棒性。本文提出的MSA-PF-GAN架构在此基础上,结合了多尺度特征融合和渐进式解码策略,以适应小样本数据集下的故

障特征提取与生成。

MSA-PF-GAN 的框架模型如图 1 所示。该网络结构接受某类别的 k 张图片 ($x_1, x_2, \dots, x_k$) 作为输入,从输入的图片中选择一张图像作为基准图像,其余的图片作为参考图像。在图像生成部分,输入的图片经过编码器后编码为特征向量 e_x ,然后经过局部融合模块将深度特征与随机向量生成融合特征,最后经过编码器生成新的图像 x_f 。在图像判别模块,输入的特征经过卷积层后再次转化为局部特征向量,然后经过多尺度注意力模块从全局和局部两个维度选取重点区域,减少多轮匹配产生

的伪影,并通过局部损失将结果返回到生成器中,从而影响生成器给明显特征的区域更高的权重,进一步提升生成图像的质量。

结合生成式对抗网络,根据特征之间的余弦相似度对局部特征进行匹配、选择以及融合。加入多尺度多种类型的注意力机制,选择关键的局部特征,提高图像生成质量并提升图像多样性。改进渐进式解码器,对生成图像进行更细粒度的控制,进一步提升生成图像的多样性。

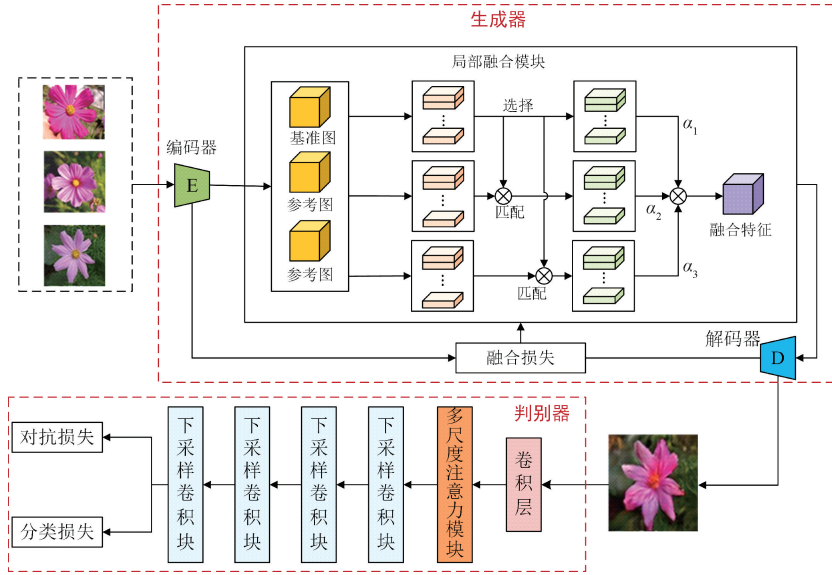


图 1 MSA-PF-GAN 的框架

Fig. 1 Framework diagram of MSA-PF-GAN

1.1 特征融合模块

图片经过编码器编码后生成深层次特征 F ,将编码后的特征分为基准特征图 ϕ_{base} 和参考特征图 f_{ref} 。从特征图 F 中随机选择一个特征图作为基准特征,将其余特征图作为参考特征 f_{ref} 。基准特征图 ϕ_{base} 作为融合过程中的主要特征图,用于对比和选择最具代表性的局部特征;参考特征图 f_{ref} 则提供了丰富的上下文信息,用于补充和增强生成图像的多样性。整个融合过程分为选择特征、匹配特征和融合特征 3 个步骤。

1) 选择特征,将选择好的基准特征划分为 $c \times h \times w$ 的维度,设定选择参数 $\eta \in (0, 1]$,控制局部特征的选择比例,从基准特征的 $h \times w$ 维度中随机选取 n 个局部特征进行替换。通过随机选择,这些局部特征从基准特征图中提取,以便在融合过程中能够代表整个图像的主要信息。

2) 匹配特征,将选定的基准特征与参考特征逐一进行余弦相似度匹配,其计算公式如式(1)所示。

$$M^{(i,j)} = g(\phi_{base}^{(i)}, f_{ref}^{(j)}) \quad (1)$$

式中: g 代表余弦相似度; $M^{(i,j)}$ 代表计算相似度结果。

3) 融合特征:通过匹配的过程,选出与基准特征最相似的参考特征,并对这些参考特征进行融合。使用随机系数向量 β 融合所有选出的参考特征得到一个融合特征,其中融合的公式如式(2)所示。

$$\phi_{fuse}^{(t)} = \beta_{base} \phi_{base}^{(t)} + \sum_{i=1, \dots, k; i \neq base} \beta_i \phi_{ref}^{(i)}(t) \quad (2)$$

式中: $t = 1, \dots, n$ 。最后经过一个卷积层和一个批归一化层的序列变换得到一个特征融合图作为输出。

1.2 多尺度注意力模块

多尺度注意力模块由两个支路组成,一个支路为注意力机制层,一个支路为池化层、注意力层和上采样层。输入的图像被多尺度提取,得到不同尺度的特征。其中一路是从通道维度选择,使用 SELayer^[16] 来增强特征图的通道间关系,其中对于注意力机制的公式为:

$$Attn(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \times V \quad (3)$$

式中: Q 、 K 、 V 分别表示对查询矩阵、键矩阵和值矩阵, 代表可训练的权重矩阵; d_k 是键向量的维度, 最后用加权平均的方式得出输入矩阵中每个元素与其他元素的相关性。另一路从空间维度选择, 使用了池化层降低图像的分辨率; 使用 SELayer 来选择图像的全局重要特征, 通过上采样层复原图像分辨率。最后将两路特征进行加和, 输入到后续的残差网络中, 其计算如式(4)所示。

$$Attn(x) = \alpha_1 Attn1 + \alpha_2 Attn2 \quad (4)$$

式中: α_1 和 α_2 表示一个随机系数向量; $Attn(x)$ 表示两路特征加和后的结果。引入注意力机制进行调整, 其结构特征图如图 2 所示。

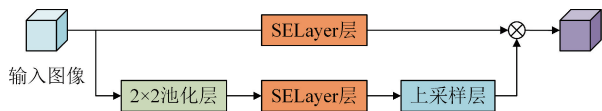


图2 多尺度注意力模块

Fig. 2 Multi-scale attention module diagram

多尺度注意力模块在齿轮箱故障诊断中能有效聚焦于故障图像中的关键区域, 抑制无关信息的干扰。通过从全局和局部两个维度选择重要特征区域, 减少了生成过程中可能出现的伪影, 从而提高生成图像的清晰度和诊断准确性。

1.3 解码器模块

借鉴于风格迁移的思想, 设计渐进式解码特征的方式, 将特征经过不同上采样处理后逐级注入到网络结构中, 与不同尺度特征融合, 低分辨率的特征在低层次被加入, 对应的高级语义将被呈现, 而相应的高分辨率的特征在高层级时加入网络, 对应图像的细节特征将在最终生成图像中展现。

解码器模块主要由卷积块、激活函数、上采样层以及卷积层构成。解码器结构如图 3 所示, 输入一个 $128 \times 8 \times 8$ 的特征图, 将经过两路支路, 一路特征经过上采样操作后输入卷积块后变为 $128 \times 16 \times 16$, 一路特征将经过上采样后也变为 $128 \times 16 \times 16$ 。然后将特征与经过卷积块支路的特征进行加和得到特征图 x_1 , x_1 再经过卷积块 2 和上采样层, 得到特征图 x_2 维度变为 $128 \times 32 \times 32$ 。再次分两路输入卷积块 3 和上采样支路获得特征图 x_3 。在本次上采样后经过一个 1×1 的卷积层, 修改通道数量, 特征加和后的特征图 x_4 变为 $64 \times 64 \times 64$ 维度, 再次将特征图输入两路支路中, 本次上采样层后依然添加 1×1 卷积层, 将特征图 x_5 的维度变为 $32 \times 128 \times 128$, 然后经过卷积块 5 输出最后的图像 x_f , 其维度变为 $3 \times 128 \times 128$ 。过程表示为:

$$x_{li} = Conv(F_{li}) \quad (i = 1, 2, 3, 4, 5) \quad (5)$$

$$x_{li} = Conv(F_{li}) \quad (i = 1, 2, 3, 4, 5) \quad (6)$$

$$x_f = x_{li} + x_{lj} \quad (7)$$

式中: F_{li} 代表每次直接输入卷积块的特征图; x_{li} 是经过卷积块后的特征; F_{lj} 是输入上采样层的特征图; x_{lj} 是经过上采样层后的特征; x_f 代表两路特征的加和; λ 是一个融合率参数。

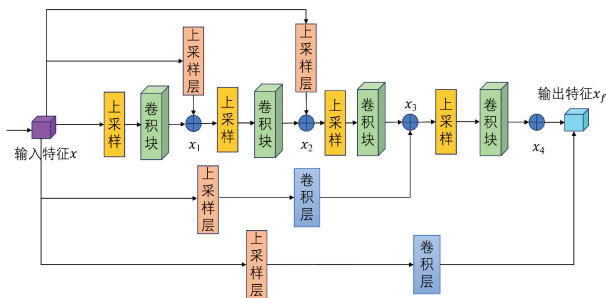


图3 解码器结构

Fig. 3 Decoder structure diagram

渐进式解码器通过逐层上采样实现特征融合, 能够逐步生成具有不同分辨率的故障特征。这一设计不仅提高了生成图像的细节表达, 还能在生成高分辨率图像时确保故障特征的准确性。

1.4 融合卷积块结构

在齿轮箱故障诊断中, 不同故障类型具有各自的特征信息, 特征融合通过选择、匹配和融合局部特征, 可以提取具有代表性的故障特征, 从而提升生成图像的诊断价值。基准特征用于对比和选择最具代表性的局部特征, 而参考特征提供了丰富的上下文信息, 以支持生成图像的多样性。

经过上采样和卷积层序列努力去还原图像的原始细节, 该过程中依然可能出现一定程度的细节损失。因此为解决以上问题, 采用渐进式的方式, 融合不同层次特征, 抵消以上的细节损失。融合卷积块的结构主要由卷积层、批量归一化层 (batch normalization, BN) 以及激活层构成。其结构如图 4 所示。其中主路的特征经过 3×3 卷积层后, 经过 BN 层进行归一化处理, 将不同层次的特征融合放在激活函数之前进行加和运算。

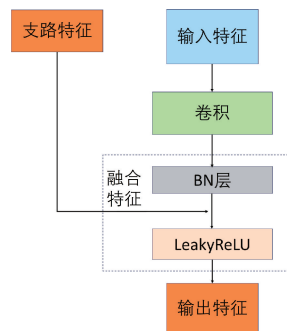


图4 卷积块结构

Fig. 4 Convolutional block structure diagram

现存的 GAN 网络架构中的非线性操作会导致高频信号的出现,这些高频信号可能包含了不应有的绝对像素坐标信息,进而导致纹理固定在屏幕坐标上而不是物体表面,形成“纹理粘附”现象。为消除这种现象,必须对网络中的信号处理方式进行重新设计。其方法是通过连续信号的方法消除高频成分,公式表达为:

$$f_{\sigma}(z) = \psi_s * \sigma(z) \quad (8)$$

$$f_{\sigma}(z) = s^2 \cdot \theta_s * \sigma(z) \quad (9)$$

$$III_s = \sum_{x \in z} \delta(x - (X + \frac{1}{2})/s) \quad (10)$$

$$F_{\sigma}(Z) = s^2 \cdot III_s \odot (\theta_s * \sigma(\theta_s * z)) \quad (11)$$

式中: ψ_s 是理想低通滤波器; s 是离散采样率; $Z[X]$ 表示离散采样特征图; $*$ 代表连续卷积; $z(x)$ 代表连续采样; III_s 表示二维的 Dirac comb; F 代表一些卷积非线性操作等; θ_s 代表理想插值滤波器 θ_s ; $f_{\sigma}(z)$ 是非线性的连续表示; $F_{\sigma}(Z)$ 代表离散的对应表示。

因此,借鉴于 StyleGAN3^[17] 提出采用连续信号表示法,并关注如何在连续域内进行操作以确保对于像素级别的平移和旋转等变性。对于卷积操作,当特征融合发生在激活函数之前时,可以在连续域中直接应用卷积,而不引入额外的高频成分,这对于维持等变性至关重要。

1.5 生成器与判别器

生成器结构由编码器、特征融合模块以及解码器构成。编码器由 1 个输入卷积块和 4 个下采样卷积块构成,每个卷积层都包含了反射填充、归一化以及激活函数等操作,输入的图像 x 经过编码器后,提取深层次特征 $F = E(X)$,特征融合模块会把得到深层次特征和一个随机系数向量 β 作为输入,生成语义对齐的融合特征,最后通过一个解码器,将得到的融合特征编码为生成图像。生成器的主要作用是通过生成逼真的合成图像扩展小样本数据集,使得故障诊断模型能够更全面地学习不同故障模式的特征。

判别器结构如图 5 所示,由注意力机制模块以及两路损失构成。首先输入图像经过 1 个空洞卷积,增加感受野;再经过多尺度注意力模块,选择重要的特征,经过 4 个残差块,并采用残差结构;最后会经过两个不同的输出分支,一路损失函数判断输入图像是真实图像或是生成图像,另一路损失用于判断输入图像的分类。判别器通过多尺度注意力模块有效地识别出合成图像中的细节差异,从而提高生成图像的真实性,确保生成样本具有良好的诊断效果。

1.6 损失函数

为了确保生成的图像具有诊断价值,本文设计了多种损失函数。对抗损失用以生成逼真的合成图像;分类损失通过类别识别确保生成图像的多样性和准确性;局

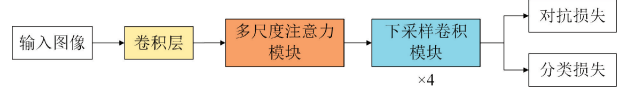


图 5 判别器结构

Fig. 5 Discriminator structure diagram

部特征损失则通过局部区域的特征匹配增强生成图像中关键故障特征的表达。通过损失函数来交替优化生成器 G 和判别器 D 。使用铰链损失^[18]来约束生成器生成的图像,使判别器无法判别真假。其计算公式为:

$$L_{adv}^D = \max(0, 1 - D(x)) + \max(0, 1 + D(\hat{x})) \quad (12)$$

$$L_{adv}^G = -D(\hat{x}) \quad (13)$$

分类损失采用 ACGAN 中的损失函数,其中应用辅助分类器将输入的图像分类到相应的类别。其计算公式为:

$$L_{cls}^D = -\log P(c(X) | X) \quad (14)$$

$$L_{cls}^G = -\log P(c(X) | \hat{X}) \quad (15)$$

式中: X 表示输入图像; $c(X)$ 表示 X 的标签。

给定一组输入图像和随机系数,引入局部特征损失,记录选择特征和参考特征的位置,将选定的特征每个位置映射到原始图像,获得简略的融合图像,具体的损失函数如式(16)所示。

$$L_{fuse} = \|\hat{x} - \sum_{i=1}^k a_i \cdot x_i\|_1 \quad (16)$$

式中: x_i 表示输入的一组图像; a 表示随机系数向量; $\hat{x}$ 表示采用加权局部特征损失约束生成的图像,最后使用下列目标函数对整个网络进行端到端的优化。

其中生成器计算公式为:

$$L_G = L_{adv}^G + \lambda_{cls}^G L_{cls}^G + \lambda_{fuse}^G L_{fuse} \quad (17)$$

判别器的计算公式为:

$$L_D = L_{adv}^D + \lambda_{cls}^D L_{cls}^D \quad (18)$$

2 实验结果

2.1 数据集介绍

行星齿轮箱图像故障诊断实验数据集 1 采用康涅狄格大学齿轮箱测量的振动信号数据^[19]。实验设备主要包括电机、电机控制器、两级行星齿轮箱和制动器。输入轴的转速由转速计测量,齿轮的振动信号由 dSPACE 系统测量,采样频率设置为 20 kHz。引入输入轴上 9 种不同的工况来收集各种故障的振动信号数据集,包括健康、缺齿、根部裂纹、剥落和尖端崩落 5 种不同严重程度。每种类型有 100 个样本,每个样本包含 3 600 个采样点的时域振动信号。

实验数据集 2 的行星齿轮箱数据集来自于东南大学

的传动系统动态模拟器^[20],行星齿轮箱故障模拟平台的两种不同情况。

行星齿轮箱工况数据采集于转速系统负载为 20 Hz-0 V 或 30 Hz-2 V,采样频率为 2 kHz。每种工况下包括 5 种故障类型,分别是齿尖崩裂、缺齿、根部裂纹、表面磨损和健康状态。采集的数据为 30 Hz-2 V 下的电机振动信号,每个故障类型有 100 个样本,每个样本包含 2 000 个时域振动采样点。利用连续小波变换,将时域信号转换为 64×64 pixels 的 RGB 三通道时频图,并通过热图形式可视化。

不同的故障形式(如缺齿、根部裂纹、表面磨损等)会在振动信号中表现出不同的频率特征。例如,缺齿故障往往表现为周期性幅值的显著波动,而表面磨损可能对应较高频率的连续波动。通过连续小波变换,这些故障特征被映射为时频图中的不同纹理和颜色分布,使得复杂的振动特征能够以图像形式可视化呈现,从而帮助深度学习模型自动学习和识别故障模式。这种故障形式到图像特征的映射关系为 MSA-PF-GAN 的故障样本生成和分类提供了有力支持。

因此,行星齿轮箱故障数据集 1 和 2 分别转换为 9 和 5 个类别的图像数据集。为验证 MSA-PF-GAN 在小样本数据增强的有效性,在每类故障类型中仅选择 6 张图像用于训练,测试集包含每类 100 张图像。这样的设置可以真实模拟实际应用中样本不足的情况,从而验证本文方法在小样本条件下的适用性。而通过采用两个独立的数据集进行验证,进一步评估了 MSA-PF-GAN 的泛化能力和稳健性,以确保模型在不同数据分布和采样条件下的表现一致性。表 1 为整个数据集数据情况。

表 1 齿轮箱时频图详细信息

Table 1 Gearbox spectrogram details

数据集	工作状态	维度	训练集	测试集
行星齿轮箱 1	健康	[3,64,64]	0	100
	缺齿	[3,64,64]	6	100
	根部裂纹	[3,64,64]	6	100
	剥落	[3,64,64]	6	100
	尖端崩落 1	[3,64,64]	6	100
	尖端崩落 2	[3,64,64]	6	100
	尖端崩落 3	[3,64,64]	6	100
	尖端崩落 4	[3,64,64]	6	100
	尖端崩落 5	[3,64,64]	6	100
行星齿轮箱 2	健康	[3,64,64]	0	100
	剥落	[3,64,64]	6	100
	缺齿	[3,64,64]	6	100
	根部裂纹	[3,64,64]	6	100
	表面磨损	[3,64,64]	6	100

2.2 实验环境设置

实验的主要参数如表 2 所示。

表 2 训练过程中的参数设置

Table 2 Parameter settings during training

参数名称	参数值
模型训练次数 (epochs)	100 000
判别器步长	1
k-shot	3
Batch size	6
选择率	0.3
对抗损失的超参数设置	1.0
分类损失的判别器超参数	$\lambda_{cls}^D = 1.0$
分类损失的生成器超参数	$\lambda_{cls}^G = 1.0$
局部融合特征损失的超参数	$\lambda_{fuse} = 0.5$

实验共计迭代次数 100 000 次,首先使用 0.000 1 的固定学习率进行 50 000 次迭代训练,然后使用线性衰减学习率进行 50 000 次迭代训练。利用梯度惩罚正则化用于增强训练过程的稳定性。最后,每种类型获得 50 个生成样本。

2.3 生成样本质量的评估

将 MSA-PF-GAN 的性能与一些主流的 GAN 模型进行了比较,包括 ILoFGAN、LoFGAN、VAEGAN^[21]、WGAN_GP^[22]、ACGAN 和 DCGAN。所有的模型都使用每种类型的 6 个故障样本训练,进行了定量和定性的分析。首先进行定量分析,实验计算了两个评价指标的来评估各种方法的生成质量:使用 Frechet 变换初始距离 (frechet inception distance, FID) 评估生成样本的多样性和相似性,使用结构相似性指数 (structural similarity index measure, SSIM) 分数评估两副图像在亮度、对比度和结构等方面的相似性。每个生成对抗模型的评估结果如表 3 所示。

表 3 各方法的 FID 分数和 SSIM 分数

Table 3 FID score and SSIM score of each method

模型	FID(↓)		SSIM(↑)	
	齿轮箱 1	齿轮箱 2	齿轮箱 1	齿轮箱 2
MSAPFGAN	62.15	116.861 4	0.496 9	0.402 3
ILoFGAN	78.00	123.010 0	0.370 0	0.309 8
WGAN_GP	104.96	170.220 0	0.340 0	0.240 0
LoFGAN	140.98	164.360 0	0.350 0	0.308 6
VAEGAN	152.79	266.140 0	0.190 0	0.110 0
DCGAN	190.42	398.550 0	0.230 0	0.260 0
ACGAN	208.75	319.960 0	0.150 0	0.240 0

从表 3 可以看出,相对于主流的 GAN 模型,MSA-PF-GAN 模型可以获得最低的 FID 分数,也获得了较好的 SSIM 分数。在行星齿轮箱 1 上,MSA-PF-GAN 方法相较于 ILoFGAN 方法 FID 下降了 15.85,SSIM 分数提升了 0.126 9;相较于 WGAN_GP 方法 FID 下降了 42.18,SSIM 分数提升了 0.156 9。在行星齿轮箱 2 上,MSA-PF-GAN 方法相较于 ILoFGAN 方法 FID 下降了 6.148 6,SSIM 分

数提升了 0.092 5;相较于 LoFGAN 方法 FID 分数下降了 47.498 6,SSIM 提升了 0.162 3。

MSA-PF-GAN 在少样本条件下的优异表现,表明该方法可以更有效地从有限的故障样本中挖掘细节特征并生成多样化的图像。这一结果突出显示了 MSA-PF-GAN 在小样本数据集上的适用性。这种方法不仅避免了因样本不足导致的过拟合问题,还在生成图像上明显优于现有的小样本图像生成方法。

此外,小样本数据集通常面临样本多样性不足的问题。MSA-PF-GAN 的渐进式解码特征融合策略,通过逐层上采样和特征融合,使得生成图像能够逐步呈现多样化的细节特征。尽管基于 MSA-PF-GAN 的图像生成方法增加了诊断流程的复杂性,但其在故障诊断的准确性、鲁棒性和泛化能力方面表现出显著优势。生成的高质量图像提升了诊断的精度,并在不同故障模式和数据稀缺的情况下保持稳定性,显示出较强的抗噪能力。同时,低 FID 和高 SSIM 分数进一步证明了方法的泛化能力,使得模型能够适应多种数据分布,从而为工业故障诊断提供了一种可靠且有效的解决方案。

针对每个数据集中典型的故障类型(例如根部裂纹、剥落、碎裂以及表面故障),MSA-PF-GAN 方法的 SSIM 分数与其他模型进行对比,具体分数如图 6 和 7 所示,左边虚线柱表示 MSA-PF-GAN 方法,中间加粗表示 IloFGAN 方法,最右侧表示 LoFGAN 方法。

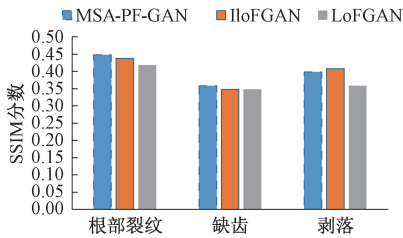


图 6 行星齿轮箱 1 故障类型的 SSIM
Fig. 6 Planetary gearbox 1 fault type SSIM

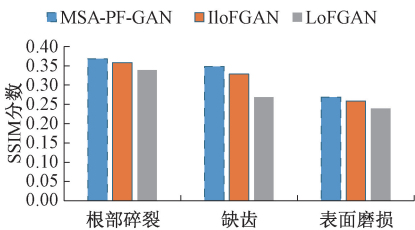


图 7 行星齿轮箱 2 故障类型的 SSIM
Fig. 7 Planetary gearbox 2 fault type SSIM

从图 7 可以看出,几类故障的 SSIM 均值高于其他模型。即使在某些故障类型上,其他模型的 SSIM 值接近 MSA-PF-GAN,但 MSA-PF-GAN 生成图像的 FID 值仍优

于 ILoFGAN。实验说明,小样本生成图像方法能在有限样本的情况下生成真实的图片,验证了多尺度注意机制在选择有用特征方面的效果,同时渐进式特征融合结构有效捕获了更多图像细节,生成了更加多样且真实的图像。

除了定量实验,还对模型进行了定性评估,比较了 MSA-PF-GAN、ILoFGAN 和 LoFGAN 在生成行星齿轮箱故障时频图方面的表现。3 种模型针对第 1 个数据集生成的图像对比如图 8 所示。从图 8 可以看出,MSA-PF-GAN 生成的图像在清晰度和多样性上与真实样本相似,表现出较好的生成效果。例如,在第 1 个行星齿轮箱数据集中,MSA-PF-GAN 相较于 ILoFGAN 和 LoFGAN 模型生成了更多的波段,增加了图像的多样性;在图像清晰度方面,MSA-PF-GAN 相较于 LoFGAN 有显著提升。在第 2 个数据集中,MSA-PF-GAN 方法生成的波段较 ILoFGAN 更清晰,而 LoFGAN 生成的图像较为模糊。整体生成效果如图 9 所示。因此,从可视化角度验证了 MSA-PF-GAN 方法在生成故障样本时频图方面的效果。

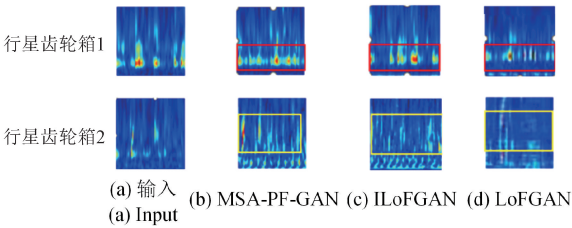


图 8 生成时频图的重点区域对比
Fig. 8 Generate a comparison chart of key areas of the time-frequency graph

2.4 消融实验

使用 MSA-PF-GAN 模型为行星齿轮箱生成图像,其目标是进行数据增强提升故障诊断的正确率。先用残差网络 ResNet 对两个行星齿轮箱数据集进行原数据集的故障诊断,然后使用 MSA-PF-GAN 模型依次为两个故障数据集生成时频图,最开始只增加少量的生成图,然后依次增加生成图像的数量,观察生成图像对于故障诊断精准确度的影响。

故障分类准确率使用的两个故障数据集,其中无故障的图像样本有 70 张,故障样本每类数量有 20 张,生成样本的故障数量会从 0 增加到 40 张图像,然后每次测试时,任意选择其中的 30 张图像作为测试样本。两个故障诊断数据集的测试结果如表 4 所示。

在行星齿轮箱 1 上,相较于原始数据集,增加 40 张故障样本时,准确率提升 35%;在行星齿轮箱 2 上,相较于原始数据集,增加 40 张故障生成样本时,准确率提升了 20%,证明 MSA-PF-GAN 小样本图像生成方法应用于

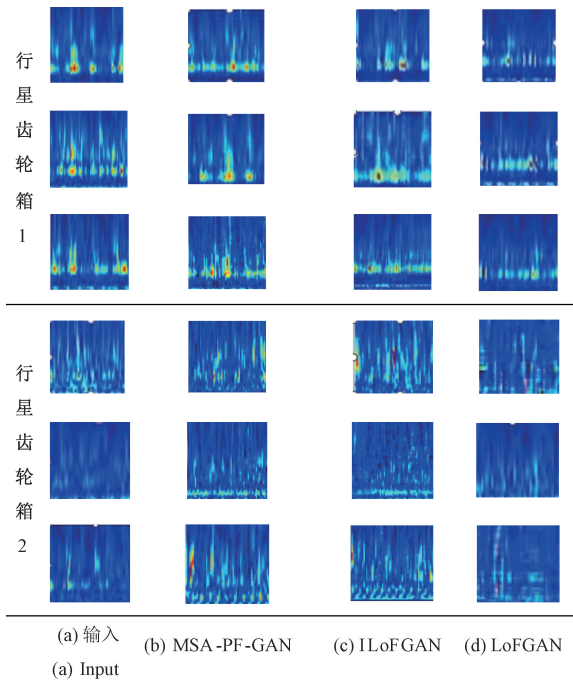


图 9 生成时频样本图像

Fig. 9 Generate time-frequency sample images

表 4 基于数据增强的故障诊断准确

Table 4 Accurate fault diagnosis based on data enhancement

添加的生成样本数量	行星齿轮箱 1	行星齿轮箱 2
0	0.46	0.55
10	0.59	0.66
20	0.65	0.71
30	0.78	0.74
40	0.81	0.75
50	0.75	0.72

数据增强可以提升故障诊断的准确率。进一步增加样本数量会引发性能下降,因此需要关注生成样本的数量和质量平衡。

考虑超参数对生成图像的影响,在行星齿轮箱 1 上将选择率从 0.1 取值到 1 探究其对于生成质量 FID 指数的影响。如图 10 所示,当选择率过低时候,融合的区域会较少,FID 分数也会略高,但是当选择率过高时,融合过程会不稳定,FID 会上升较多。

考虑到融合率也会影响到图像的生成质量,在行星齿轮箱 2 上测试不同融合率对于生成质量 FID 分数的影响。如图 11 所示,当融合率为 1 时,获得最好的 FID 分数,意味着网络在处理上采样后的信息与经过卷积块提取的特征信息时,给予了两者同等重要的地位。这代表了网络能同时有效地利用高分辨率的细节信息以及经过卷积压缩提炼的关键特征信息来合成高质量的样本,表

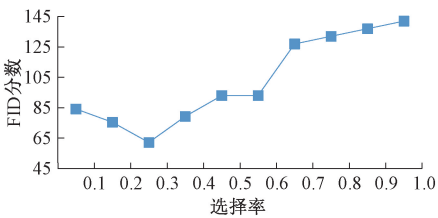


图 10 不同选择率下的 FID 分数

Fig. 10 FID scores under different selection rates

明此时模型能够在保留全局结构的同时兼顾局部细节,实现了一种平衡。这种平衡可能是通过优化信息融合机制,有效解决了梯度消失或爆炸等问题,使得从底层到高层的特征都能得到充分的学习和表达。

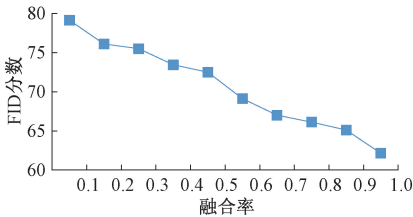


图 11 不同融合率下的 FID 分数

Fig. 11 FID scores at different fusion rates

为探索不同尺度特征的融合顺序对于齿轮箱图像生成效果的影响,分别测试特征融合放在激活函数之前和特征融合放在激活函数之后的 FID 分数和 SSIM 分数,选用数据集 1 进行实验,其测试的结果如表 5 所示。

表 5 不同特征融合顺序测试的 FID 和 SSIM 结果

Table 5 FID and SSIM results of different feature fusion order tests

顺序	齿轮箱数据集 1	
	FID	SSIM
激活函数之前	62.15	0.496 9
激活函数之后	82.62	0.482 9

从表 5 可以看出,在激活函数之前整合不同层次的特征,能够有效提升图像质量,相对于在激活函数之后融合 FID 分数降低了 20.47,同时增强 SSIM 指标。这意味着网络能更高效地操纵和保留细节特征,避免了过度依赖绝对像素位置信息,从而促进了生成图像多样性的提升。进一步证实了通过调整非线性运算的执行顺序来优化模型的策略是合理的。

通过定量实验探索 MSA-PF-GAN 方法各个模块对于齿轮箱图像最终生成效果的影响,主要研究多尺度注意力模块以及渐进式解码器模块两个部分对于生成结果影响的消融实验。采用齿轮箱 1 的数据集,实验设置依然为 3-shot。实验结果如表 6 所示。

表 6 齿轮箱 1 上的不同功能模块消融实验
Table 6 Ablation experiment of different functional modules on gearbox 1

设置	FID(↓)	SSIM(↑)
无多尺度注意力模块	80.16	0.493 6
无渐进式解码器模块	73.37	0.488 1
MSA-PF-GAN	62.15	0.496 9

定量实验包含 FID 和 SSIM 两个量化的指标。可以发现,MSA-PF-GAN 方法相对于没有多尺度注意力模块的方法,FID 分数下降了 17.99,SSIM 分数提升了 0.003 3,相对于没有渐进式解码器结构的方法,FID 分数下降了 11.22,SSIM 分数提升了 0.008 8。因此不论是当多尺度注意力模块缺失或者是渐进式解码器缺失时,MSA-PF-GAN 模型的生成效果都会有所下降,这进一步说明以上两个模块对于生成图像都有正向促进作用。

3 结 论

本文围绕行星齿轮箱故障诊断的小样本数据问题,提出了一种基于多尺度渐进式特征融合的生成对抗网络,通过渐进式解码器逐层整合不同尺度的特征信息,结合多尺度注意力机制精确提取关键区域,从而在生成高质量故障图像的同时显著提升故障诊断的精度与鲁棒性。这一方法突破了传统生成对抗网络在小样本数据下的瓶颈,解决了生成样本语义错位、伪影问题等难点,为小样本学习场景提供了一种新型的特征表达和优化框架。实验结果表明,该方法在多个数据集上实现了生成图像质量和诊断准确率的双重提升,体现了较强的泛化能力和实用价值。本文通过创新性的特征融合和多尺度建模技术,使得模型能够在低样本条件下有效捕捉复杂故障特征,显著提升诊断系统的智能化水平,为未来工业设备智能监测提供了理论基础和技术参考。未来工作将进一步优化网络结构,提升生成速度与样本分布合理性,同时探索该方法在其他复杂机械系统中的通用性,以推动智能故障诊断技术的广泛应用。

参考文献

[1] 刘洋. 行星齿轮箱故障诊断技术的研究进展[J]. 内燃机与配件, 2018(18): 154-155.
LIU Y. Research progress on fault diagnosis technology of planetary gearboxes[J]. Internal Combustion Engine & Parts, 2018 (18): 154-155.
[2] 陈伟,王友仁,孙灿飞,等. 行星齿轮箱耦合调制振动信号分离与故障诊断[J]. 航空动力学报, 2018, 33(5): 1112-1120.
CHEN W, WANG Y R, SUN C F, et al. Separation and

fault diagnosis of coupled modulation vibration signals in planetary gearboxes [J]. Journal of Aerospace Power, 2018, 33(5): 1112-1120.
[3] 朱静,邓艾东,李晶,等. 基于改进 EEMD 和声发射技术的行星齿轮箱故障诊断研究[J]. 噪声与振动控制, 2018, 38(S2): 657-660.
ZHU J, DENG AI D, LI J, et al. Research on fault diagnosis of planetary gearbox based on improved EEMD and acoustic emission technology [J]. Noise and Vibration Control, 2018, 38(S2): 657-660.
[4] 林丽,邓春,经昊达,等. 基于油液在线监测的齿轮箱磨损趋势分析与研究[J]. 材料导报, 2018, 32(18): 3230-3234.
LIN L, DENG CH, JING H D, et al. Analysis and research on gearbox wear trends based on online oil monitoring [J]. Materials Herald, 2018, 32 (18): 3230-3234.
[5] 胡莒庆,陈徽鹏,程哲,等. 基于经验模态分解和深度卷积神经网络的行星齿轮箱故障诊断方法[J]. 机械工程学报, 2019, 55(7): 9-18.
HU N Q, CHEN H P, CHENG ZH, et al. Planetary gearbox fault diagnosis method based on empirical mode decomposition and deep convolutional neural network[J]. Chinese Journal of Mechanical Engineering, 2019, 55(7): 9-18.
[6] 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. Planetary gearbox fault diagnosis based on feature fusion and ResNet[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(5): 213-222.
[7] 王辉,徐佳文,严如强. 基于多尺度注意力深度强化学习网络的行星齿轮箱智能诊断方法[J]. 机械工程学报, 2022, 58(11): 133-142.
WANG H, XU J W, YAN R Q. Intelligent diagnosis method of planetary gearbox based on multi-scale attention deep reinforcement learning network [J]. Chinese Journal of Mechanical Engineering, 2022, 58(11): 133-142.
[8] 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. Gearbox fault diagnosis method using temporal attention boundary enhanced prototype network under small samples [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(2): 90-98.

- [9] ANTONIOU A, STORKEY A, EDWARDS H. Data augmentation generative adversarial networks[C]. ArXiv preprint arXiv:1711.04340, 2017.
- [10] CLOUÂTRE L, DEMERS M. Figr: Few-shot image generation with reptile [J]. ArXiv preprint arXiv: 1901.02199, 2019.
- [11] ANTONIOU A. Data augmentation generative adversarial networks[J]. ArXiv preprint arXiv:1711.04340, 2017.
- [12] HONG Y, NIU L, ZHANG J F, et al. MatchingGAN: Matching-based few-shot image generation [C]. 2020 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2020: 1-6.
- [13] HONG Y, NIU L, ZHANG J F, et al. F2GAN: Fusing-and-filling gan for few-shot image generation [C]. Proceedings of the 28th ACM International Conference on Multimedia. 2020: 2535-2543.
- [14] GU ZH, LI W B, HUO J, et al. LoFGAN: Fusing local representations for few-shot image generation [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 8463-8471.
- [15] CHEN M, SHAO H, DOU H, et al. Data augmentation and intelligent fault diagnosis of planetary gearbox using ILoFGAN under extremely limited samples [J]. IEEE Transactions on Reliability, 2022, 72(3): 1029-1037.
- [16] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [17] NICHOL A, ACHIAM J, SCHULMAN J. On first-order meta-learning algorithms [J]. ArXiv preprint arXiv: 1803.02999, 2018.
- [18] ODENA A, OLAH C, SHLENS J. Conditional image synthesis with auxiliary classifier gans [C]. International Conference on Machine Learning. PMLR, 2017: 2642-2651.
- [19] CAO P, ZHANG SH L, TANG J. Preprocessing-free gear fault diagnosis using small datasets with deep

convolutional neural network-based transfer learning[J]. IEEE Access, 2018, 6: 26241-26253.

- [20] SHAO S Y, MCALEER S, YAN R Q, et al. Highly accurate machine fault diagnosis using deep transfer learning [J]. IEEE Transactions on Industrial Informatics, 2018, 15(4): 2446-2455.
- [21] LARSEN A B L, SØNDERBY S K, LAROCHELLE H, et al. Autoencoding beyond pixels using a learned similarity metric [C]. International Conference on Machine Learning. PMLR, 2016: 1558-1566.
- [22] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks [C]. International Conference on Machine Learning. PMLR, 2017: 214-223.

作者简介



高文超(通信作者), 2013 年于中国矿业大学(北京)获得博士学位, 现为中国矿业大学(北京)副教授, 主要研究方向为人工智能及工业故障诊断。

E-mail: gaowe@cumb.edu.cn

Gao Wenchao (Corresponding author)

received her Ph. D. degree from China University of Mining and Technology (Beijing) in 2013. She is currently an associate professor at China University of Mining and Technology (Beijing). Her main research interests include artificial intelligence and industrial fault diagnosis.



陈一帆, 2022 年于中国矿业大学(北京)获得学士学位, 现为中国矿业大学(北京)硕士研究生, 主要研究方向为图像识别处理及齿轮箱故障诊断。

E-mail: cyf08132024@163.com

Chen Yifan received his B. Sc. degree from China University of Mining and Technology (Beijing) in 2022. He is currently a M. Sc. candidate at China University of Mining and Technology (Beijing). His main research interests include image recognition processing and gearbox fault diagnosis.