

基于GAN和少样本学习的电机故障诊断方法研究

邹琪红

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年2月20日; 录用日期: 2025年3月27日; 发布日期: 2025年4月8日

摘要

随着工业自动化的进步, 电机故障诊断已成为保证工业设备正常运行的关键技术。传统的故障诊断方法往往依赖于大量标注数据进行训练, 但实际应用中, 电机的故障种类繁多且故障数据较为稀少, 数据不足成了故障诊断的一大瓶颈, 为了克服这一挑战, 本文提出了基于对抗生成网络的故障诊断方法, 旨在解决数据问题, 提高模型性能。本文引入了对抗生成网络, 通过增强数据来缓解数据不足的问题。具体来说, 我们先将振动数据进行频谱化, 然后作为对抗生成网络的输入, 通过生成器生成更多的数据, 用于扩充数据集, 生成的数据不仅具有较好的时频特征, 还保留故障的多样性, 从而更好地提升小样本学习模型的泛化能力。经过增强后的数据再经过LKA (Large Kernel Attention) 模块进行特征提取, 最后经过一个全局分支一个局部分支处理后进行分类。此外为了进一步提升诊断精度和训练效率, 我们结合了KL散度和Wasserstein距离, 提出并采用了动态权重调整策略和学习率调整策略, 使得训练过程更加稳定, 并加速了优化过程。本文在公开数据集CWRU上进行了大量实验, 结果表明了我们所提模型的有效性。

关键词

电机故障, 诊断模型, 深度学习, GAN

Research on Motor Fault Diagnosis Method Based on GAN and Few-Shot Learning

Qihong Zou

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Feb. 20th, 2025; accepted: Mar. 27th, 2025; published: Apr. 8th, 2025

Abstract

With the advancement of industrial automation, motor fault diagnosis has become a key technology

文章引用: 邹琪红. 基于 GAN 和少样本学习的电机故障诊断方法研究[J]. 软件工程与应用, 2025, 14(2): 176-188.
DOI: 10.12677/sea.2025.142017

to ensure the normal operation of industrial equipment. Traditional fault diagnosis methods often rely on large amounts of labeled data for training. However, in practical applications, motor faults are varied, and fault data is relatively scarce, making data insufficiency a major bottleneck for fault diagnosis. To address this challenge, this paper proposes a fault diagnosis method based on Generative Adversarial Networks (GANs) to solve the data issue and improve model performance. In this study, we introduce GANs to enhance data and alleviate the problem of data insufficiency. Specifically, we first convert vibration data into spectrograms, which are then used as input to the GAN. The generator generates additional data to expand the dataset. The generated data not only possesses good time-frequency features but also retains the diversity of faults, thereby improving the generalization ability of the few-shot learning model. The enhanced data is then processed through the LKA (Large Kernel Attention) module for feature extraction. Finally, after processing by a global branch and a local branch, the data is classified. Furthermore, to further improve diagnostic accuracy and training efficiency, we combine KL divergence and Wasserstein distance, and propose dynamic weight adjustment and learning rate adjustment strategies. These strategies make the training process more stable and accelerate optimization. Extensive experiments on the public CWRU dataset demonstrate the effectiveness of the proposed model.

Keywords

Motor Fault, Diagnosis Model, Deep Learning, GAN

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

工业生产中设备或系统的监测与维护是一个重要课题。有统计数据显示,在尚未计算故障本身带来的生产损失下,各类企业的维护成本通常就达到了总支出的15%~60%,冶金、采矿企业更是高达50%~60% [1]。国内外大量机组故障案例表明一旦机组发生故障,轻则影响机组的寿命和整个生产线的稳定性,重则造成巨大的经济损失和严重的安全事故[2] [3]。

在众多电机故障中,其中轴承故障通常占机械故障的40%~50% [4]。轴承的健康状态可以直接反映电机的运行状态,而振动信号是反映轴承故障特征的最直接和常用的物理量,通过对振动信号进行分析可以精确识别故障类型、位置和严重程度[5]。文章[6]总结了基于振动信号和声学测量的轴承故障诊断技术,指出振动信号因其高灵敏度和丰富的特征信息而成为诊断滚动轴承故障的首选。进而轴承振动数据信号成为电机故障诊断的理想数据来源。但由于数据较少,所以对于数据的处理和有效特征的提取显得格外重要。

传统的信号特征处理与机器学习技术在处理非线性复杂信号方面有瓶颈,而深度学习以其强大的特征处理能力和建模能力,在预测、图像处理、故障识别等诸多领域都表现出巨大的潜力。本文将采用生成对抗网络(Generative adversarial networks, GAN) [7]对数据进行加强。GAN在数据生成方面具有独特的效应,它可以弥补数据不足的缺点,为解决故障诊断识别这类小样本问题提供了新思路[8]。然而传统的GAN存在模式崩塌(Mode Collapse) [9]等问题,且对复杂的时序数据的特征提取和生成能力有限。此外,Transformer [10]在处理高维信号特征时计算成本较高,所需样本要足够大,所以它在小样本问题中的泛化能力[11]也不足。

受上述工作启发,本文提出了一种基于GAN和少样本学习的网络模型,并首次将该模型引入故障识

别领域,不同于传统的生成器判别器结构,本文引入 Kullback-Leibler (KL)散度来更新生成器,引入 Wasserstein GAN (WGAN)来更新判别器,与此同时通过设置 dropout 层来缓解过拟合问题。此外,我们提出了一种动态权重调整策略。同时,我们引入了学习率调整机制,专为生成对抗网络(GAN)在电机故障诊断中的应用而设计。本文的主要贡献如下:

- 1) 在该领域引入了一种新的复合函数:结合 Wasserstein 距离和 KL 散度,利用 Wasserstein 距离捕捉生成数据与真实数据分布的全局差异,同时通过 KL 散度优化局部细节,使生成对抗网络(GAN)在电机故障诊断中的数据建模更加全面和精确。
- 2) 引入了动态权重调整策略:在训练过程中,根据生成样本与真实样本的差异,自适应地调整 KL 散度和 Wasserstein 距离的权重,实现了生成器和判别器的动态平衡,提升了模型的收敛速度。
- 3) 引入动态学习率调整机制:在训练过程中,根据 KL 散度和 Wasserstein 距离的权重动态变化来动态调整学习率,使模型训练更稳定,精度更高。
- 4) 在公开数据集 CWRU 上进行了大量消融实验,证明我们所提方法的优越性。

2. 国内外研究现状

近些年来深度学习的发展,逐渐应用到了故障诊断领域。深度学习方法相比于传统的方法,解决了一些时间和成本上的问题,但故障数据不足仍然是一大挑战,对抗生成网络旨在通过两个神经网络生成接近真实分布的数据分布,与真实分布一起作为整体模型的输入,从而弥补数据不足的缺点。同时少样本学习是一种元学习方法,也可用于上述诊断问题。

在故障识别等领域,生成对抗网络(Generative adversarial networks, GAN)的引入为数据稀缺问题提供了新的解决方案。例如, Mirza 等人[12]提出了将 GAN 用于振动信号生成的模型,通过引入条件 GAN,实现了对不同故障类别振动信号的条件生成,以扩展数据分布的多样性。此外, Zhao 等人[13]将 GAN 与卷积神经网络结合,设计了一种生成增强数据并进行端到端故障分类的模型。该模型有效地缓解了数据不平衡问题,并提高了少样本情况下的诊断性能。

在数据特征学习方面, Zhang 等人[14]提出了一种新的滤波式变换器(FTGAN),并利用神经网络对离散傅里叶变换进行近似,训练过程中学习从时域到频域的通用映射。与此同时, Gulrajani 等人[15]改进了 Wasserstein GAN (WGAN),通过加入梯度惩罚项,解决了传统 GAN 训练不稳定的问题,并实现了在强噪声干扰条件下的可靠数据生成。尽管 GAN 在故障诊断与识别中的应用优于传统的方法,但是其生成器生成的样本质量仍存在挑战,为确保生成数据更接近真实分布并加强模型运行的稳定性,许多研究者开始探索生成器的判别标准和优化版本, J Song 等人[16]提出了桥接 f-GAN 和 WGAN 的理论框架,并在此框架下,探讨了 KL 散度与 Wasserstein 距离的结合,同时提出了改进的生成对抗目标函数,证明了 KL-WGAN 在生成质量和数据多样性上有显著效果。这种框架目前并未应用到故障诊断识别领域,且其主要关注 KL 散度优化,依赖于判别器的准确性,同时可能受到权重估计偏差的影响。目前的少样本学习模型虽已取得一定的成效,但都需要大规模数据集,并在图卷积神经网络或 Transformer 为基础的模型框架上进行预训练,这显然不适用于数据稀少的轴承数据集。

3. 整体设计

3.1. 模型框架

由于在故障诊断中,不同故障模式的频率特征相较于时序特征会更明显,而噪声影响会减弱,同时可初步平滑掉一些不重要的特征,所以本文先利用短时傅里叶变换(Short-Time Fourier Transform, STFT)将时序数据转为时频图,作为所提 GAN 模块的输入,这样有利于生成器学习到除数据分布外的时频相关

性, 并保留原始数据的周期性、频率成分和时频特性。为了弥补数据较少的缺陷, 本文提出了结合动态权重调整策略和学习率调整机制的 KL_WGAN 模型, 数据经过 GAN 加强之后, 采用少样本学习方法进行分类, 得到支持集和查询集, 为从支持集查询集中提取特征, 采用了基于 LKA (Large Kernel Attention) 的多重特征提取模块。由于我们需要测量支持集和查询集之间的相关性, 本文采取了 Transformer 框架来处理全局信息以及相关性, 采用 Vu 等人[17]所提出的局部马氏距离来处理支持集的局部特征。最后采用注意力机制构成的分类模块, 将全局信息和局部信息结合起来实现故障分类。本文所提模型的整体框架如图 1 所示。

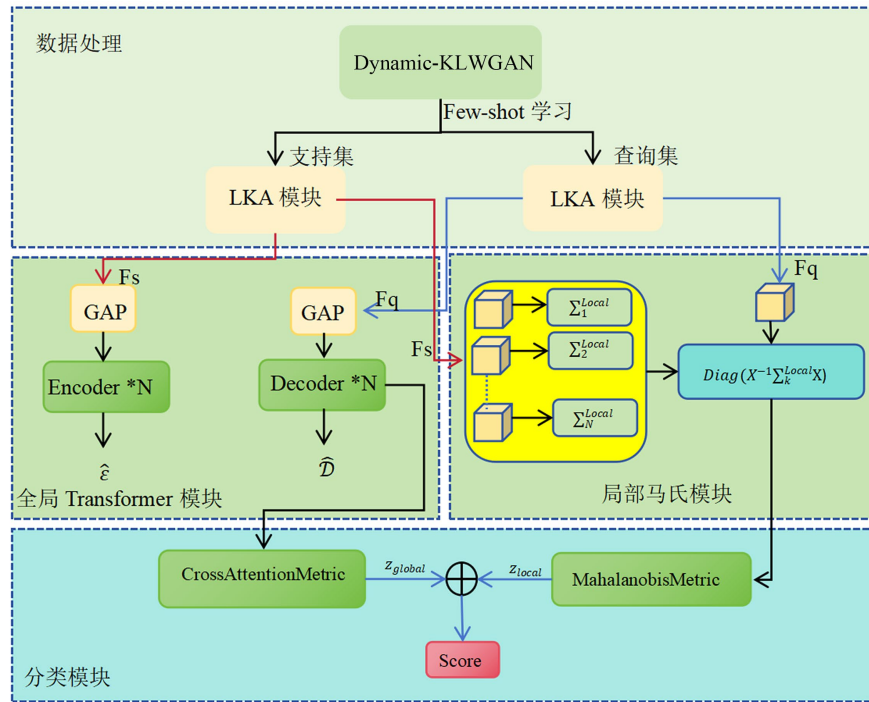


Figure 1. The overall architecture of Model
图 1. 模型整体框架图

3.2. 数据处理

在本文中, 我们采用少样本学习与对抗生成训练结合的方法来尽可能弥补无法收集大量故障数据集的缺陷。由于轴承故障数据集难以大量收集, 我们采用 Meta Learning 中的 Few-shot 方法, 并根据一些文献研究, 最终采用了已在少样本学习中成功应用的“episode 训练机制”。Few-shot 训练集中包含了很多类别, 在训练阶段, 该种机制将在训练集中随机选取 C 个类别, 每个类别包含 K 个样本(共 CK 个数据), 来构成一个 meta-task, 用作支持集(support set) $S = \left\{ \left(X_i^s, Y_i^s \right) \right\}_{i=1}^{n_s}$ (其中, $n_s = C * K$) 的输入, 同时在每类中随机选取 q 个样本用于构建查询集(Query set)。

$$Q = \left\{ \left(X_j^q, Y_j^q \right) \right\}_{j=1}^{n_q} \quad (1)$$

其中, $n_q = q * K$ 。每次 episode 都会采样得到不同的 meta-task, 这种机制使得模型可以学习到不同 meta-task 中的共性, 也能对不在支持集中的元任务进行较好地分类。少样本学习的基本思想就是学习一个相似性函数, 通过比较查询集和支持集之间的相似度来判断查询集中某一样本属于支持集的哪一类。

3.3. 特征提取模块

本文先对原始的频谱图像进行多层卷积处理(包括 Conv2d、Maxpooling、LayerNorm 和 GELU 激活函数), 其中早期层捕获高频成分, 更深层次捕获全局信息。为找到某一时间或频率上的特征, 采用了不同卷积核大小的卷积操作。

对于特征提取模块, 我们以大核注意(Large Kernel Attention, LKA)模块为基础, 这个模块将大核卷积分解为深度卷积、扩张深度卷积以及逐点卷积, 然后再将他们进行组合, 用于捕获长距离关系。由于 LKA 属于固定配置类型的方法, 而故障数据是复杂的、非线性的, 需要同时处理其局部和全局信息, 所以在本文中设置了不同配置的 LKA, 分别用于提取局部特征和全局特征, 其中不同的模块卷积核和扩张率基于小卷积核、小扩张率以及大卷积核、大扩张率进行选择, 最后通过残差连接来保留原数据的细节信息, 获取短距离关系。为避免特征融合之后部分特征过强或者过弱, 通过逐点卷积和深度卷积来进行特征融合和输出。

3.4. Dynamic-KLWGAN

生成对抗网络(GAN, Generative Adversarial Network)的核心思想由 Goodfellow 等人[12]在 2014 年提出, 主要采用生成器和判别器对数据进行对抗生成训练, 通过上述的对抗优化, 可以得到逼真的数据。尽管 GAN 在对抗生成训练中取得巨大成功, 但是其训练需要生成器和判别器相对稳定, 即判别器的过强生成器过弱会导致对抗训练失败, 在[14]中作者提出了将使用 Wasserstein 距离代替传统的 JS 散度, 改善了训练不稳定性。但由于故障信号本身的多样性, WGAN 也无法涵盖真实数据的多样性, 所以本文引入了 Kullback-Leibler (KL)散度与 WGAN 结合来缓解模式崩塌问题。同时引入了动态权重调整策略和学习率调整机制, 所提的 Dynamic-KLWGAN 的模型图如图 2 所示。

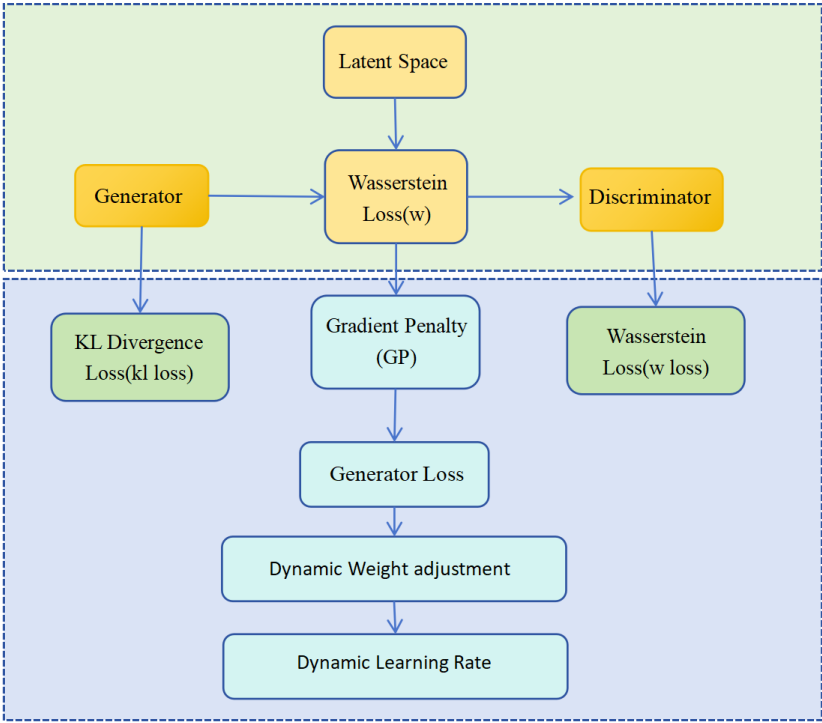


Figure 2. The architecture of Dynamic-KLWGAN Model
图 2. Dynamic-KLWGAN 模型框架图

A. GAN

GAN 的对抗训练可以概括为极小化 - 极大化目标函数:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_r} [\log D(x)] + \mathbb{E}_{z \sim P_g} [\log(1 - D(G(z)))] \quad (2)$$

其中, P_r 是真实数据分布, P_g 是生成数据分布, $D(x)$ 表示判别器输出真实样本的概率, $G(z)$ 是生成器生成的样本。

B. KL 散度

本文引入 KL 散度作为辅助损失, 与判别器损失一起优化生成器参数, KL 散度是一个非对称的度量, 用于衡量两个概率分布 $P_r(x)$ 和 $P_g(x)$ 之间的差异, 即真实分布 $P_r(x)$ 相较于近似分布 $P_g(x)$ 的额外的信息代价, 其连续形式的数学定义如下:

$$D_{KL}(P_r \parallel P_g) = \int P_r(x) \log \frac{P_r(x)}{P_g(x)} dx \quad (3)$$

其中, $P_r(x)$ 为真实数据分布, $P_g(x)$ 为生成数据分布, KL 散度用于指导生成器的优化, 从而使 $P_g(x)$ 更加接近于 $P_r(x)$ 。受 f-GAN [16] 中框架启发, 我们将 KL 散度用于 WGAN 中, Wasserstein 距离更注重全局分布, 而 KL 散度对高概率区域的差异更为敏感, 因此可以弥补 Wasserstein 距离在局部分布细节上的不足。KL 部分的损失函数可以表示为:

$$\mathcal{L}_{KL} = \mathbb{E}_{x \sim P_r} [\log P_g(x)] - \mathbb{E}_{x \sim P_r} [\log P_r(x)] \quad (4)$$

其中, $\mathbb{E}_{x \sim P_r} [\log P_g(x)]$ 是固定常数, 所以该部分的目标是最小化 $-\mathbb{E}_{x \sim P_r} [\log P_r(x)]$ 部分, 即可最大化生成样本的概率密度。

C. Wasserstein 距离

为优化 GAN 的目标函数, 解决梯度消失等问题, 采用了 Wasserstein 距离, 它表示将一个分布中的质量“搬运”到另一个分布所需的最小代价:

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} \mathbb{E}_{(x, y) \sim \gamma} [\|x - y\|] \quad (5)$$

其中, $\Pi(P_r, P_g)$ 是所有可能的联合分布的集合, P_r, P_g 分别为两个边缘概率, $\|x - y\|$ 表示将质量从 x 移动到 y 所需的代价。Wasserstein 距离是连续的, 即使真实分布和生成分布没有重叠(即分布之间的差距较大), 也可以提供有意义的梯度, 从而保证生成器的有效更新。WGAN 的目标函数基于 Kantorovich-Rubinstein 对偶性, 可以通过以下形式表达:

$$W(P_r, P_g) = \sup_{\|f\|_L \leq 1} \mathbb{E}_{x \sim P_r} [f(x)] - \mathbb{E}_{x \sim P_g} [f(x)] \quad (6)$$

其中, f 是一个 1-Lipschitz 函数。在训练过程中, 判别器会被设计为一个接近 1-Lipschitz 的网络。由于 1-Lipschitz 可能引入优化问题, 且权重裁剪的方法会导致梯度更新不稳定, 当裁剪范围较小时, 判别器可能难以充分学习到数据分布, 所以本文采用梯度惩罚(Gradient Penalty, GP)来实现 Lipschitz 条件。判别器目标函数更新为:

$$\mathcal{L}_D = \mathbb{E}_{x \sim P_g} [D(x)] - \mathbb{E}_{x \sim P_r} [D(x)] + \lambda_{gp} \mathbb{E}_{\hat{x} \sim P_{\hat{x}}} \left[\left(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1 \right)^2 \right] \quad (7)$$

其中, $\lambda_{gp} \mathbb{E}_{\hat{x} \sim P_{\hat{x}}} \left[\left(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1 \right)^2 \right]$ 是梯度惩罚项, 用于强制判别器的梯度范数归为 1, $P_{\hat{x}}$ 是真实数据和生成数据之间的随机插值分布, $\hat{x}$ 如下所示:

$$\hat{x} = \epsilon x_r + (1 - \epsilon) x_g, \quad \epsilon \sim \text{Uniform}(0, 1) \quad (8)$$

D. Dynamic-KLWGAN

KL-WGAN 是一种结合 KL 散度和 Wasserstein 距离的生成对抗网络。传统的 WGAN 主要优化 Wasserstein 距离, 这里我们为了更好的捕捉到数据分布之间的差异, 更好的衡量真实分布与生成分布的差异, 引用了 f-GAN 框架[16], 引入 KL 散度作为辅助损失, 不同的是, 本研究在 KL-WGAN 的基础上提出了动态权重调整策略和动态学习率机制, 接下来详细介绍该策略和机制。

对于权重更新规则, 本文根据 KL 散度和 Wasserstein 距离的变化对权重比例进行处理, KL 散度权重更新公式如下所示:

$$w_{\text{KL}} = \begin{cases} \min(w_{\text{KL}} + \Delta_{\text{KL}}, w_{\text{KL}, \max}) & \text{if } \text{KL} > \text{KL}_{\text{threshold}} \\ \max(w_{\text{KL}} - \Delta_{\text{KL}}, w_{\text{KL}, \min}) & \text{if } \text{KL} \leq \text{KL}_{\text{threshold}} \end{cases} \quad (9)$$

其中, Δ_{KL} 为 KL 权重的步长, 控制更新速度, $w_{\text{KL}, \max}$ 和 $w_{\text{KL}, \min}$ 为 KL 散度的最大最小权重, $\text{KL}_{\text{threshold}}$ 为 KL 散度的阈值。

Wasserstein 距离权重的更新公式如下所示:

$$w_{\text{W}} = \begin{cases} \min(w_{\text{W}} + \Delta_{\text{W}}, w_{\text{W}, \max}) & \text{if } W > W_{\text{threshold}} \\ \max(w_{\text{W}} - \Delta_{\text{W}}, w_{\text{W}, \min}) & \text{if } W \leq W_{\text{threshold}} \end{cases} \quad (10)$$

其中, Δ_{W} 为 Wasserstein 距离权重的步长, 控制更新速度, $w_{\text{W}, \max}$ 和 $w_{\text{W}, \min}$ 为 Wasserstein 距离的最大最小权重, $W_{\text{threshold}}$ 为 Wasserstein 距离的阈值。本文根据 KL 散度和 Wasserstein 距离的变化去更新权重, 更新规则是当 KL 散度较大时, 增加其权重, 强化生成器对 KL 散度的优化能力, 当 KL 散度较大时, 减小其权重, 降低对 KL 的约束, 平衡 Wasserstein 距离的优化。当 Wasserstein 距离较大时说明生成器与判别器之间的分布差异较大, 需要加强 Wasserstein 权重, 距离较小时, 降低其权重以免过拟合。

为进一步提升模型的训练效率和稳定性, 我们结合动态调整策略引入了学习率(Learning Rate, LR)调整机制, 将学习率的上下限设置为 $[\text{lr}_{\min}, \text{lr}_{\max}]$, 并采用 Pytorch 中的 LambdaLR, 它的核心思想是通过自定义的 $\text{lr}_{\text{lambda}}(t)$ 函数, 即根据动态权重的变化比例来动态更新学习率, 当权重增大时, 学习率略微减少, 从而减少振动, 当权重减小时, 学习率略微增加, 从而加速收敛, 提升训练速率。本研究先结合动态权重调整策略设计 $\text{lr}_{\text{lambda}}(t)$ 函数如下所示:

$$\text{lr}_{\text{lambda}}(t) = \begin{cases} 1 + \alpha & \text{if } w_{\text{kl}} > \theta_{\text{kl}} \text{ and } w_{\text{w}} > \theta_{\text{w}} \\ 1 - \beta & \text{if } w_{\text{kl}} < \theta_{\text{kl}} \text{ or } w_{\text{w}} < \theta_{\text{w}} \\ 1 & \text{otherwise} \end{cases} \quad (11)$$

简化为:

$$\text{lr}_{\text{lambda}}(t) = (1 + \delta \cdot \Delta w) \quad (12)$$

其中, α 、 β 和 δ 为调整因子, w_{kl} 和 w_{w} 分别为 KL 散度和 Wasserstein 距离的权重, 为增强模型平衡力, 设置了两个阈值, θ_{kl} 和 θ_{w} 分别为 w_{kl} 和 w_{w} 的阈值, 通常根据经验值设置为 0.5, 当 w_{kl} 和 w_{w} 同时较大时, 此时模型已经逐步学习到主要特征, 可以适当增加学习率, 加快收敛, 当某一项权重过大时, 说明模型的某一特征学习不足, 需要降低学习率, 从而稳定优化。实际学习率更新公式为:

$$\eta_t = \eta_{t-1} \cdot (1 + \delta \cdot \Delta w) \quad (13)$$

其中, Δw 表示权重变化比例, α 表示调整因子。从而判别器损失函数可以由上述公式更新为:

$$\mathcal{L}_D = -\omega_w \cdot \mathbb{E}_{x \sim P_r} [D(x)] + \mathbb{E}_{x \sim P_g} [D(x)] + \lambda_{gp} \cdot \mathcal{L}_{gp} \quad (14)$$

其中, ω_w 是 Wasserstein 距离的动态权重, λ_{gp} 是梯度惩罚项的权重, $\mathcal{L}_{gp}$ 为梯度惩罚项。生成器损失包括 KL 散度损失, 可以由以下表示出来:

$$\mathcal{L}_G = -\mathbb{E}_{x \sim P_g} [D(x)] + \omega_{kl} \cdot \mathcal{L}_{KL} \quad (15)$$

其中, ω_{kl} 为动态 KL 权重, $\mathcal{L}_{KL}$ 为生成分布于真实分布的 KL 散度。所以我们的总损失函数定义为:

$$\mathcal{L}_{Total} = \mathcal{L}_D + \mathcal{L}_G \quad (16)$$

这里我们首次将 KL_WGAN 框架用于电机故障识别领域, 在这个模块训练中, 我们首先初始化生成器和判别器的权重、学习率优化器和优化器。再在每个训练步计算 Wasserstein 距离和 KL 散度, 接着使用上述提到的动态权重调整策略和学习率调整机制来更新权重和学习率, 并计算 KL 散度的损失函数和 Wasserstein 损失函数, 用于更新和优化生成器和判别器, 循环训练 5 次, 最后生成近似真实分布的生成数据, 与真实数据一起作为 Fewshot 模型的输入, 从而有效缓解故障诊断中的小样本问题, 提高模型对复杂故障模式的识别能力, 进而为一些工业场景提供更为准确、鲁棒的诊断解决方案。同时, 动态权重策略和学习率机制增强了原始模型模型的适应性, 使其更适合实际应用场景中的多样化需求。

4. 实验设计与验证

4.1. 数据集与参数设置

为验证本文所提方法的有效性, 采用著名的 CWRU 公开数据集对模型进行训练和评估:

CWRU 数据集是关于轴承故障的数据集, 轴承主要由内圈(Inner Race)、外圈(Outer Race)和滚动体(Rolling Elements)组成, 这三个部分是导致故障的主要原因, 本文将其按照故障类型、转速、具体故障类型进行分类, 具体见表 1, 故障类型包括内圈故障、外圈故障、滚动体故障和正常状态, 转速分为 1730 rpm、1750 rpm、1772 rpm、1797 rpm, 具体故障类型有与该部分裂纹相关的三个小故障 Ball、InnerRace、OuterRace。本文采用滑动窗口的方法对信号数据进行切片处理, 采用数据集前半部分作为训练集, 后半部分作为测试集, 训练样本采用 2048 点滑动窗口机制, 滑动步长是 80。测试样本采用相同步长的滑动窗口集步长为序列长度。为了验证模型的有效性, 本文随机选取了 750 个测试样本。为保证训练过程的随机性和公平性, 以及增加模型的泛化能力, 我们在实验中对数据进行了打乱并在测试过程中加入了噪声。

Table 1. CWRU Dataset description

表 1. CWRU 数据集信息

故障类型	故障直径	故障标签
Ball	0.007	1
	0.014	2
	0.021	3
InnerRace	0.007	4
	0.014	5
	0.021	6
OuterRace	0.007	7
	0.014	8
	0.021	9
Health		0

模型设置：在所提的特征提取模块，我们设置了 3 个 LKA 模块，卷积维度为 64，每个模块卷积核大小和扩张率分别为(3, 3)、(5, 3)、(7, 3)，特征图通道数设置为 4096。全局变换器(Global Transformer)中的编码器 - 解码器数量选择为 $N=10$ 。超参数 α_1 和 α_2 分别是局部分支和全局分支在最终输出中的比例，根据消融实验最终选择 $\alpha_1 = 0.7$ ， $\alpha_2 = 0.3$ 。根据消融实验和以往经验，GAN 网络采用 Adam 优化器，其 dropout 层参数设置为 0.5，初始学习率为 0.0002，动态学习率范围为(0.0001, 0.001)，KL 散度权重 ω_{kl} 动态范围为(0.1, 0.5)，Wasserstein 距离权重 ω_w 动态范围为(0.5, 2.0)，梯度惩罚权重 λ_{gp} 设置为 10。

4.2. 评价指标与环境

本研究中，为更好地评价模型性能并与其他方法公平比较，我们采用了准确率、召回率(Recall)作为评价指标。召回率用于消融研究，它是正确预测为正样本的数量(TP)与实际正样本数量(TP + FN)的比值，在本研究中，首先通过模型的输出分数来预测查询样本的类别，再与实际标签进行比较，分别对每个类别计算召回率，最后计算得到召回率的平均值，定义如下：

$$\text{Recall} = \frac{1}{C} \sum_{c=1}^C \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c} \quad (17)$$

其中， C 是类别总数，TP (True Positive, 真正类)表示被正确预测为正类的样本数量。FN (False Negative, 假负类)表示为实际为正类但被错误预测为负类的样本数量。

在 Few-shot 任务中，本文会计算每次模型测试中每个查询集样本的预测结果，对于每个批次的准确率的计算，我们根据模型输出的分数，采用最大值索引来选择预测的类别，并与实际标签进行对比，从而计算得到模型的准确率，计算公式如下：

$$\text{Accuracy} = \frac{\sum_{i=1}^N \mathbb{I}(\text{argmax}(\text{scores}_i) = \text{target}_i)}{N} \quad (18)$$

其中， N 为查询样本总数， $\mathbb{I}$ 为指示函数，当预测的类别与真实标签一致时返回 1，否则返回 0。

为了训练本文提出的模型，我们的模型采用 Adam 优化器来优化训练，学习率按照经验设置为 0.001，每十次训练减半，降至 0.000001 之后学习率稳定不变。本实验中模型训练 epoch 设置为 100，GPU 使用 RTX 2080 Ti x2 (22GB)，CPU 使用 9 vCPU Intel(R) Xeon(R) Platinum 8255C CPU @ 2.50GHz。

4.3. 实验结果与分析

本节展示所提模型与其他不含对抗生成网络的少样本模型的对比结果。其中度量列 Euclidean 为欧氏距离，cov 为协方差度量(cov, covariance metric)，CA 为交叉注意力度量(cross-attention metric)，EMD 为地球移动距离度量(Earth'mover distance metric, EMD)。(1) Shen 等人[18]直接使用原始振动信号和频谱信号作为输入，使用神经网络来提取特征，并使用原始网络学习各类故障数据的特征，采用欧氏距离作为故障分类的度量。(2) Hao 等人[19]提出了一种新颖的交叉注意力网络来处理未知类别问题，并提出了传导推理算法以减轻低数据问题。(3) Li 等人[20]定义了一种新的深度协方差指标，用于衡量查询样本与新概念之间的分布一致性。(4) Chang 等人[21]解决了利用上下文信息来寻找危险区域的对应关系的手段。(5) X. Wang 等人[22]针对少样本学习的核心问题提出了查询集 - 支持集 Transformer 模型(Query-Support TransFormer model, QSFormer)，并提出了一种新的跨尺度互动特征提取器(Cross-scale Interactive Feature Extractor, CIFE)。(6) Manh-Hung Vu 等人[17]就小样本分类问题引用了 Transformer 全局分支和改进的局部马氏距离分支。本文引用其中全局分支和局部分支的框架，并采用无监督学习。本文中前六种方法的

实验结果均来自[17]。

表 2 所展示的是所提模型与其他模型的对比结果，首先，可以看出在数据集数量较少时，仅采用有监督学习和欧氏距离度量的少样本模型表现较差，30 样本时仅 61.13%，300 样本时仅 78.77%。covaMNet [20]和 MF-Net [21]采用协方差度量和仅采用有监督学习的少样本学习模型在样本极少时表现较差，样本较少时有大幅度提升，可以得出提取并融合数据的全局特征和局部特征的模型可以有效提升准确率，表现为 30 样本时准确率分别为 72.17%和 83.21%，600 样本时准确率分别为 96.51%和 97.51%。QS-Former [22]和 Fewshot-cross-Mah [17]采用了 Transformer 结构来获取全局信息，可以看出采用马氏距离获取数据局部特征的方法不论在样本少还是多的情况下都优于 EMD。在 CWRU 数据集上，本文所提出的模型优于其他六种模型，相比较[17]的模型，本文所提模型在样本 30, 600, 19800 时分别提升了 17.20%、0.09%、0.04%。这证明采用生成对抗网络对数据进行增强，从而有效扩充训练集，并有效帮助小样本学习模型学习到数据之间的相似性和差异性，进而提升了分类准确性。

表 3 所展示的是本文所提模型与[17]中所示模型的运行时间对比。可以得出增加对抗生成网络的模型可以以更少的训练周期达到较高的性能，减少了训练时间，同时由于数据先经过 STFT 成频谱图的形式可以使得数据特征更为明显，模型训练时能够快速聚焦于重要特征，进而降低训练的计算量。并且可以利用生成对抗网络的并行生成合成数据的能力，有效利用多核处理器，提高训练效率。

Table 2. Comparison of Proposed model accuracy (%) with other methods

表 2. 所提模型的准确率(%)与其他方法对比

方法	度量方式		训练样本数		
	度量	数据增强模型	30	600	19800
ProtoNet [18]	Euclidean	unenhan	61.13	96.61	97.62
CrossAttentionNetwork [19]	CA	unenhan	72.17	96.83	99.29
CovaMNet [20]	cov	unenhan	79.73	98.91	99.62
MF-Net [21]	cov	unenhan	83.21	99.58	99.63
QS-Former [22]	CA + EMD	unenhan	79.88	99.06	99.63
Fewshot-cross-Mah [17]	Cross + Mah	unenhan	84.87	99.78	99.89
本文所提模型	Cross + Mah	KL-WGAN*	99.47	99.87	99.93

Table 3. Comparison of Proposed model training time (min) with other methods

表 3. 所提模型的运行时间(min)与其他方法对比

模型	训练样本		
	30	600	19800
Fewshot-cross-Mah [17]	203.45	190.63	220.71
本文所提模型	182.85	191.57	207.92

4.4. 消融实验

为了清晰展示所提模型各个模块的改进有效性，我们进行了消融实验。本节评估 KL-WGAN 变种模型的性能包括运行效率，以评估相关组件的作用。基准模型最初基于 GAN 构造，该网络由生成器和判别器组成，并引入 Wasserstein 距离作为他们之间的度量。我们将改进后的 WGAN 结合[17]中所提的 Transformer 全局模块和马氏距离局部模块构成新的故障诊断模型。具体工作如下：本文引入 KL 散度优化生成器，并提出动态权重调整策略以及学习率调整机制。在数据有限的情况下提高了准确率和召回率

并降低了运行时间。

表 4 中案例 1 代表仅使用 GAN 模块对数据进行处理, 案例 2 为 KL-WGAN, 引入了 KL 散度和 Wasserstein 距离优化生成器和判别器。案例 3 在 KL-WGAN 的基础上引入了 KL 动态权重调整策略。案例 4 在 KL-WGAN 的基础上提出了动态权重调整策略和学习率调整机制, 用于优化和稳定对抗生成网络。仅使用基线模型, 在数据量有限的情况下, 其结果准确率最差。在案例 2 中, 使用 KL-WGAN 对数据进行加强, 在样本为 30 的情况下, 模型的准确率和召回率达到了 99.06%和 99.20%。同时由表 5 可知其运行时间在样本增多时比基线模型时长更短。在案例 3 中仅引入动态权重调整策略, 由表 4 和表 6 可以看出样本数为 600 时, 其召回率和准确率有小幅提升, 样本数为 30 时, 准确率相较于基线模型分别提升了 16.57%。相比较 KL-WGAN 在准确率和召回率上都有所降低, 说明仅对 KL 权重进行简单的动态调整只能小幅度提升模型的分类精度, 但由表 6 可知提升了运行效率。此外, 如果在 KL-WGAN 上引入动态权重调整策略和学习率调整机制, 如案例 4 所示, 本文提出的模型相较于前面三个都有提升, 相比基线模型, 准确率在 30 和 600 样本时分别提升了 17.19%、0.08%。总体而言, 结果表明我们所提所有模块的组合是合理的。并在 30 样本和 600 样本时达到了最高(99.46%和 99.86%), 运行时间减少近 18 分钟。

Table 4. Ablation study of the accuracy and recall rate (%) of the proposed Dynamic-KLWGAN variant models

表 4. 所提 Dynamic-KLWGAN 各变种模型的准确率和召回率(%)的消融实验

	Baseline	KL-WGAN	Adaptive	Dynamic	样本数(准确率)	
					30	600
Δ1	√				84.87	99.78
Δ2	√	√			99.06	99.73
Δ3	√	√	√		98.93	99.86
Δ4	√	√		√	99.46	99.86

Table 5. Ablation study of the running time (min) of each variant model of the proposed Dynamic-KLWGAN

表 5. 所提 Dynamic-KLWGAN 各变种模型的运行时间(min)的消融实验

方法	训练样本数	
	30	600
Fewshot-cross-Mah [17]	203.45	190.63
KL-WGAN	197.91	197.40
Adaptive-klwgan	179.23	188.12
Dynamic-KLWGAN	182.85	191.57

Table 6. Ablation study of recall rate (%) of the proposed Dynamic-KLWGAN variant models

表 6. 所提 Dynamic-KLWGAN 各变种模型的召回率(%)的消融实验

方法	样本数(召回率)	
	30	600
KL-WGAN	99.20	99.87
Adaptive-KLWGAN	98.93	99.67
Dynamic-KLWGAN	99.13	99.93

5. 总结

在本研究中, 我们引入了一种新的网络模型 Dynamic-KLWGAN, 旨在解决电机故障诊断中面临少样

本学习和数据不平衡问题。这个方法主要用于数据增强部分。原始振动数据经过频谱化之后,作为对抗生成网络的真实数据,有效扩充了训练集,帮助模型更好地学习到故障特征。首先,Dynamic-KLWGAN模型通过引入动态权重调整策略,根据KL散度和Wasserstein距离的变化动态调整它们的权重,从而使生成器和判别器的优化更加平衡和高效。为了进一步提升模型的训练效率和稳定性,模型采用了学习率更新机制,在训练过程中根据权重动态的变化调整学习率,这两种方法的结合更好地平衡KL散度和Wasserstein距离的贡献,避免传统模型中梯度不稳定的问题,从而有效提高故障诊断模型的训练稳定性、少样本时的分类精度以及模型收敛速度。增强后的数据进一步由多个LKA模块进行特征提取,再经过Transformer模块捕捉到故障数据之间潜在的长距离依赖关系,马氏距离则捕捉局部特征之间的相似性进行故障分类。我们在公共数据集CWRU上进行了大量实验,结果表明,我们所提模块在数据不平衡的故障领域上性能更好,优于其他一些方法。

在未来的工作中,本文所提模型有几个可改进之处,首先是频谱化方面,可以选择一些其他类型的时频分析方法,从而提高模型对故障特征的表征能力。此外,马氏距离受特征分布的影响较大,可以适当加入自适应机制。同时,还可将此模型进一步扩展到其他故障领域。

参考文献

- [1] Mobley, R.K. (2002) An Introduction to Predictive Maintenance. Elsevier.
- [2] 杨建东, 赵琨, 李玲, 等. 浅析俄罗斯萨扬-舒申斯克水电站 7 号和 9 号机组事故原因[J]. 水力发电学报, 2011, 30(4): 226-234.
- [3] 魏炳漳, 姬长青. 高速大容量发电电动机转子的稳定性——惠州抽水蓄能电站 1 号机转子磁极事故的教训[J]. 水力发电, 2010, 36(9): 57-60.
- [4] Caesarendra, W., Widodo, A. and Yang, B. (2010) Application of Relevance Vector Machine and Logistic Regression for Machine Degradation Assessment. *Mechanical Systems and Signal Processing*, **24**, 1161-1171. <https://doi.org/10.1016/j.ymssp.2009.10.011>
- [5] Concari, C., Franceschini, G., Tassoni, C. and Toscani, A. (2013) Validation of a Faulted Rotor Induction Machine Model with an Insightful Geometrical Interpretation of Physical Quantities. *IEEE Transactions on Industrial Electronics*, **60**, 4074-4083. <https://doi.org/10.1109/tie.2012.2228142>
- [6] Ferracuti, F., Giantomassi, A. and Longhi, S. (2013) MSPCA with KDE Thresholding to Support QC in Electrical Motors Production Line. *IFAC Proceedings Volumes*, **46**, 1542-1547. <https://doi.org/10.3182/20130619-3-ru-3018.00216>
- [7] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014) Generative Adversarial Networks. *Communications of the ACM*, **63**, 139-144. <https://doi.org/10.1145/3422622>
- [8] Pan, S.J. and Yang, Q. (2010) A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, **22**, 1345-1359. <https://doi.org/10.1109/tkde.2009.191>
- [9] Arjovsky, M. and Bottou, L. (2017) Towards Principled Methods for Training Generative Adversarial Networks. arXiv: 1701.04862. <https://doi.org/10.48550/arXiv.1701.04862>
- [10] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., et al. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [11] Wu, J., Zhao, Z., Sun, C., Yan, R. and Chen, X. (2020) Few-Shot Transfer Learning for Intelligent Fault Diagnosis of Machine. *Measurement*, **166**, Article 108202. <https://doi.org/10.1016/j.measurement.2020.108202>
- [12] Mirza, M. and Osindero, S. (2014) Conditional Generative Adversarial Nets. arXiv: 1411.1784. <https://doi.org/10.48550/arXiv.1411.1784>
- [13] Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P. and Gao, R.X. (2019) Deep Learning and Its Applications to Machine Health Monitoring. *Mechanical Systems and Signal Processing*, **115**, 213-237. <https://doi.org/10.1016/j.ymssp.2018.05.050>
- [14] Wang, H., Li, P., Lang, X., Tao, D., Ma, J. and Li, X. (2023) FTGAN: A Novel Gan-Based Data Augmentation Method Coupled Time—Frequency Domain for Imbalanced Bearing Fault Diagnosis. *IEEE Transactions on Instrumentation and Measurement*, **72**, 1-14. <https://doi.org/10.1109/tim.2023.3234095>
- [15] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. and Courville, A. (2017) Improved Training of Wasserstein GANs.

-
- Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 5769-5779.
- [16] Song, J. and Ermon, S. (2020) Bridging the Gap between f-GANs and Wasserstein GANs. *Proceedings of the 37th International Conference on Machine Learning*, Virtual, 13-18 July 2020, 9078-9087.
 - [17] Vu, M., Nguyen, V., Tran, T., Pham, V. and Lo, M. (2024) Few-Shot Bearing Fault Diagnosis via Ensembling Transformer-Based Model with Mahalanobis Distance Metric Learning from Multiscale Features. *IEEE Transactions on Instrumentation and Measurement*, **73**, 1-18. <https://doi.org/10.1109/tim.2024.3381270>
 - [18] Shen, H., Zhao, D., Wang, L. and Liu, Q. (2023) Bearing Fault Diagnosis Based on Prototypical Network. *International Conference on Mechatronics Engineering and Artificial Intelligence (MEAI 2022)*, Changsha, 11-13 November 2022, 125960D. <https://doi.org/10.1117/12.2671906>
 - [19] Hou, R., Chang, H., Ma, B., *et al.* (2019) Cross Attention Network for Few-Shot Classification. *Advances in Neural Information Processing Systems*, 32.
 - [20] Li, W., Xu, J., Huo, J., Wang, L., Gao, Y. and Luo, J. (2019) Distribution Consistency Based Covariance Metric Networks for Few-Shot Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 8642-8649. <https://doi.org/10.1609/aaai.v33i01.33018642>
 - [21] Chang, J. and Chen, Y. (2018) Pyramid Stereo Matching Network. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 5410-5418. <https://doi.org/10.1109/cvpr.2018.00567>
 - [22] Wang, X., Wang, X., Jiang, B. and Luo, B. (2023) Few-Shot Learning Meets Transformer: Unified Query-Support Transformers for Few-Shot Classification. *IEEE Transactions on Circuits and Systems for Video Technology*, **33**, 7789-7802. <https://doi.org/10.1109/tcsvt.2023.3282777>