

DOI: 10.3901/JME.2023.12.215

时变转速下基于双阈值注意力生成对抗网络和小样本的转子-轴承系统故障诊断*

邵海东¹ 李伟¹ 刘翊² 杨斌³

- (1. 湖南大学机械与运载工程学院 长沙 410082;
2. 国家先进轨道交通装备创新中心 株洲 412000;
3. 株洲中车时代电气股份有限公司 株洲 412001)

摘要: 可用故障数据的匮乏给时变转速下转子-轴承系统的端到端故障诊断带来严重挑战,生成对抗网络为解决小样本故障诊断问题提供新思路,但其仍存在梯度消失、全局关联特征学习能力较弱和训练效率较低等缺点。因此,提出一种双阈值注意力生成对抗网络,用于生成高质量的红外热成像图片,以解决时变转速下转子-轴承系统的小样本故障诊断难题。首先,结合 Wasserstein 距离和梯度惩罚设计新型对抗损失函数,避免训练过程中的梯度消失。其次,构建注意力嵌入的生成对抗网络以指导学习红外热成像图片的全局热力关联特征。最后,开发双阈值训练机制进一步提高生成样本质量和训练效率。将所提方法用于分析转子-轴承系统的实测红外热成像图片,结果表明,所提方法能辅助准确诊断时变转速及小样本下的不同故障模式,性能优于目前常用的生成对抗网络方法。

关键词: 双阈值注意力生成对抗网络; 故障诊断; 时变转速; 小样本红外热成像; 转子-轴承系统

中图分类号: TH17

Fault Diagnosis of Rotor-bearing System under Time-varying Speeds by Using Dual-threshold Attention-embedded GAN and Small Samples

SHAO Haidong¹ LI Wei¹ LIU Yi² YANG Bin³

- (1. College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082;
2. National Rail Transit Advanced Equipment Innovation Center, Zhuzhou 412000;
3. Zhuzhou CRRC Times Electric Co., Ltd., Zhuzhou 412001)

Abstract: End-to-end fault diagnosis of rotor-bearing system under time-varying speeds using a few samples is challenging. Despite generative adversarial network (GAN) provides a way to solve the problem of small-sample fault diagnosis, it still has some limitations, such as gradient vanishing, weak extraction of global correlation features, and low training efficiency. Therefore, a dual-threshold attention-embedded GAN is proposed for generating high-quality infrared thermal (IRT) images to solve small-sample fault diagnosis of rotor-bearing system under time-varying speeds. First, Wasserstein distance and gradient penalty are combined to design the new adversarial loss function to avoid gradient vanishing. Second, attention-embedded GAN is constructed to guide learn global thermal-correlation features of the IRT images. Finally, dual-threshold training mechanism is developed to further improve the generation quality and training efficiency. The proposed method is used to analyze the collected small IRT images of a rotor-bearing system, and the results show that the proposed method can accurately diagnosis different fault modes using small samples under time-varying speeds, which is superior to other popular GANs.

* 国家自然科学基金(51905160)、湖南省自然科学基金优秀青年科学基金(2021JJ20017)和国家重点研发计划(2020YFB1712100)资助项目。
20220810收到初稿, 20230215收到修改稿

Key words: dual-threshold attention-embedded GAN; fault diagnosis; time-varying speeds; small infrared thermal images; rotor-bearing system

0 前言

旋转机械作为现代工业的基石,在电力、化工、矿冶等生产制造领域应用广泛。转子-轴承系统是旋转机械的关键组件,其故障可能会导致严重的经济损失和人员伤亡。为保障安全高效的工业生产,有效诊断转子-轴承系统的健康状态以提供合理的维护策略成为目前研究的热点^[1-2]。

随着人工智能的高速发展,端到端的深度学习技术进一步提升了旋转机械故障诊断的智能化程度。在近五年来,基于卷积神经网络^[3](Convolutional neural network, CNN)、深度置信网络^[4](Deep belief network, DBN)和堆叠自动编码器^[5](Stacked auto-encoder, SAE)等深度学习模型的故障诊断研究卓有成效。然而,在工业实际中,通常难以获取充足可用的故障数据来有效训练深度学习故障诊断模型以达到高精度^[6]。因此,小样本情景下的故障诊断研究更具现实意义。

生成对抗网络(Generative adversarial network, GAN)从数据生成角度为小样本故障诊断问题提供了新的解决方案^[7]。2019年,严如强等^[8]构造了一种辅助分类 GAN(Auxiliary classifier generative adversarial network, ACGAN)生成感应电机的振动信号;2020年,朱忠奎等^[9]利用 GAN 对故障数据进行重构,对轴承、转子和齿轮箱进行故障检测;2021年,肖雄等^[10]将轴承振动信号转换为二维灰度图,并通过 GAN 进行数据增强;李巍华等^[11]设计了 K-means 聚类的卷积 GAN,实现了高精度的小样本轴承故障诊断;2022年,姜洪开等^[12]提出了一种基于 pull-away 函数的 GAN,用于轴承不平衡数据生成及其故障诊断。然而,上述基于 GAN 的小样本故障诊断研究大多基于定速条件下采集的振动信号开展,存在一定局限:首先,振动传感器需要牢固地安装在机械设备表面,在长期监测过程中可能会一定程度上影响监测对象的结构表面^[13];其次,在某些空间受限的特定情况下,振动传感器不得不安装在远离关键部件的位置,数据采集过程容易受长距离路径传播干扰^[14];最后,为了满足某些实际生产要求,通常需要改变机械设备的转速,这极易

致使振动信号呈现更强的非平稳性,进一步增加故障特征提取和故障模式诊断的难度^[15]。

根据红外辐射理论,物体表面通常存在热力分布场。旋转机械故障而引起的过度摩擦会产生异常的高温信息,因此可通过红外热成像技术获取机械运行过程中的热力分布信息,以实现故障诊断。相比于振动监测,红外热成像技术具有无需接触、覆盖面广和稳定性好等优点^[16]。近年来,基于红外热成像和深度学习的旋转机械故障诊断取得了一定的进展。2018年,JANSSENS 等^[17]将 CNN 应用于热成像图像识别,以检测滚子轴承的故障状态;2019年,JIA 等^[18]利用 CNN 和红外热成像对转子系统进行诊断;2020年,LI 等^[19]提出了基于 CNN 和热成像的转子-轴承系统故障特征学习和智能诊断方法;2021年,CHOUDHARY 等^[20]将热成像图片用于训练改进 CNN 模型以提升轴承的故障诊断精度;2022年,MAO 等^[21]构造了一种用于齿轮箱红外热成像和振动信号融合故障诊断的领域自适应 CNN 模型。

然而,上述研究依旧建立在充足可用的训练样本基础上,考虑到目前 GAN 和红外热成像在故障诊断领域的结合应用尚处于起步阶段,因此,利用 GAN 生成更多红外热成像图片以解决时变转速下小样本故障诊断问题值得进一步探索。为了实现高质量红外热成像图片的高效生成,亟需解决传统 GAN 的以下问题:① 由于对抗训练的特殊性,传统 GAN 通常很难保障训练的稳定性,易出现梯度消失现象;② 传统 GAN 难以准确学习红外热成像图片中的全局热力的关联特征;③ 为保障生成样本的质量,传统 GAN 通常需要设置冗余的训练迭代次数,导致了训练的低效率和低适应性。

综上,提出一种双阈值注意力 GAN,用于生成高质量的红外热成像图片,以解决时变转速下转子-轴承系统的小样本故障诊断问题。首先,结合 Wasserstein 距离和梯度惩罚设计新型对抗损失函数,以稳定训练过程,避免梯度消失。其次,构造注意力嵌入的 GAN 模型,指导学习红外热成像图片中的全局热力关联特征。最后,开发一种双阈值训练机制,进一步提高热成像图片生成质量和训练效率。利用不同故障状态下转子-轴承系统的红外热成像图片进行分析对比,结果验证了所提方法在应

对时变转速下小样本故障诊断任务的有效性和优越性。

1 传统 GAN 与注意力模块简介

1.1 传统 GAN

GAN 的思想来源于博弈论, 其由判别器(Discriminator)和生成器(Generator)组成双网络结构。将随机噪声 z 输入生成器中, 通过对真实样本的学习, 输入噪声经多层上采样被映射为与真实样本相似的生成样本。判别器输入的是真实样本 x 或者生成样本 $G(z)$, 经过多层的特征提取后输出概率 $D(x)$, 代表输入样本属于真实样本的概率大小。在对抗性训练过程中, 生成器逐渐学习真实样本的特征分布, 而判别器也提高了对辨别真实样本和生成样本的能力。传统 GAN 的对抗损失函数如下

$$\min_G \max_D V(G, D) = E_{z \sim P_z(z)} [\lg(1 - D(G(z)))] + E_{x \sim P_r(x)} [\lg(D(x))] \quad (1)$$

式中, $P_{r(x)}$ 和 $P_{z(z)}$ 分别为真实样本分布和高斯分布。

在优化过程中, 判别器 D 的优化目标为概率值 $D(x)$ 趋向于 1, 而概率值 $D(G(z))$ 趋向于 0; 生成器 G 的优化目标为概率值 $D(G(z))$ 趋向于 1。因此, GAN 的对抗训练过程为最大最小优化过程。

此外, 式(1)通过数学运算可转换为詹森-香农(Jensen-shannon, JS)散度, 以衡量真实样本分布和生成样本分布之间的距离^[22]。因此, 实际计算的对抗损失函数为

$$\min_G \max_D V(G, D) = 2JS(P_{r(x)} \| P_{z(z)}) - 2\lg 2 \quad (2)$$

式中, $2JS(P_{r(x)} \| P_{z(z)})$ 为 $P_{r(x)}$ 与 $P_{z(z)}$ 的 JS 散度运算。

1.2 注意力模块

注意力模块以人类视觉的关注机制进行直观解释, 实质上是对数据特征进行加权计算。自注意力作为注意力模块的一员, 可用于计算图像内部不同区域之间的相互依赖关系^[23]。因此, 可在 GAN 网络中嵌入自注意力模块来指导学习图像的全局关联特征, 如图 1 所示。

在图 1 中, 假设上一层卷积运算后得到的特征图 x 的总像素个数为 n 。首先, 使用 f 和 g 两个 1×1 卷积核对特征图 x 特征提取, 得到特征张量 $f(x)$ 和 $g(x)$ 以计算注意力特征图 β , $\beta_{i,j}$ 的定义为

$$\beta_{i,j} = \frac{\exp(s_{ij})}{\sum_{i=1}^n \exp(s_{ij})} \quad s_{ij} = f(x_i)^T g(x_j) \quad (3)$$

式中, s_{ij} 代表卷积特征图 x 中第 i 和第 j 个像素点之间的相互联系; 经过 softmax 函数概率归一化计算后, $\beta_{i,j}$ 表示卷积特征图元素 x_i 和 x_j 之间的相互联系所需要受关注程度的权重。由此可知, 注意力特征图 β 能够考虑卷积特征图 x 的全局信息。

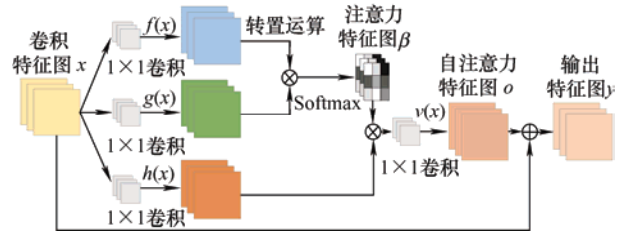


图 1 自注意力模块

随后, 将注意力特征图 β 与特征张量 $h(x)$ 进行点乘, 点乘后得到的张量经 1×1 卷积核 v 特征提取得到自注意力特征图 o 。随后将 o 与原始卷积特征图 x 相融合即可获取经自注意力模块处理后的输出特征图 y , 其计算过程如下

$$y_j = \gamma o_j + x_j \quad o_j = v\left(\sum_{i=1}^n \beta_{j,i} \times h(x_i)\right) \quad (4)$$

式中, $o_j \in o = (o_1, o_2, \dots, o_j, \dots, o_n)$; γ 为尺度参数, 将 γ 初始化为 0, 让自注意力模块由近而远的逐步学习全局信息。

2 所提方法

所提方法主要分为三部分: 新型对抗损失函数设计、注意力嵌入的 GAN 模型构造和双阈值训练机制开发。

2.1 新型对抗损失函数设计

由于 JS 散度的离散特性, 在优化过程中容易出现梯度消失的现象^[22]。相较于 JS 散度, Wasserstein 距离能更加平滑地反应两个分布之间的差异。因此, 基于 Wasserstein 距离设计的对抗损失函数可以较有效稳定训练过程^[24], Wasserstein 距离定义如式(5)所示。

$$W(P_{r(x)}, P_{g(y)}) = \inf_{\gamma \sim \prod(P_{r(x)}, P_{g(y)})} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (5)$$

式中, x 和 y 分别为真实样本和生成样本; $\prod(P_{r(x)}, P_{g(y)})$ 代表真实样本和生成样本的联合分布, 其边界为 s 和 $P_{g(y)}$; \inf 为联合分布下确界;

$E_{(x,y) \sim \gamma}[\|x - y\|]$ 为真实样本和生成样本的期望距离。

但在基于 Wasserstein 距离的 GAN(Wasserstein GAN, WGAN)中, 对判别器进行权重裁剪可能会导致梯度爆炸或者难以获得所需的函数映射^[25]。梯度惩罚能够有效地替代权重裁剪, 其定义如式(6)所示

$$gp = \lambda \times E_{\tilde{x} \sim P_{\tilde{x}}}[(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1)^2] \quad (6)$$

式中, $\varepsilon \sim \text{Uniform}[0,1]$, \tilde{x} 为 x 和 y 之间的均匀采样, 其为 $\tilde{x} = \varepsilon x + (1 - \varepsilon)y$; $\nabla_{\tilde{x}} D(\tilde{x})$ 为判别器对 \tilde{x} 输出结果的梯度; $\|\nabla_{\tilde{x}} D(\tilde{x})\|_2$ 为 $\nabla_{\tilde{x}} D(\tilde{x})$ 的 L2 范数。最终, 设计的新型对抗损失函数如式(7)所示。

$$L(G, D) = -E_{x \sim P_r(x)}[D(x)] + E_{y \sim P_g(y)}[D(y)] + \lambda \times E_{\tilde{x} \sim P_{\tilde{x}}}[(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1)^2] \quad (7)$$

式中, λ 为梯度惩罚系数。

2.2 注意力嵌入的 GAN 模型构造

由于卷积核感受野尺寸的限制, 通常需要多层卷积才能捕获到远距离像素之间的关联特征, 这极大地增加了模型参数和计算复杂度^[12]。此外, 由于卷积侧重于局部特征的提取, 在多层卷积后所提取到的关联信息也容易被稀释^[26]。在机械设备的红外热成像图片中, 分布在不同部位的热力像素值表征着相同的故障状态。考虑到上述特点, 在基于二维卷积层的判别器和生成器中分别嵌入自注意力模块, 以指导学习红外热成像图片中的全局热力关联特征。注意力嵌入的 GAN 结构如图 2 所示。

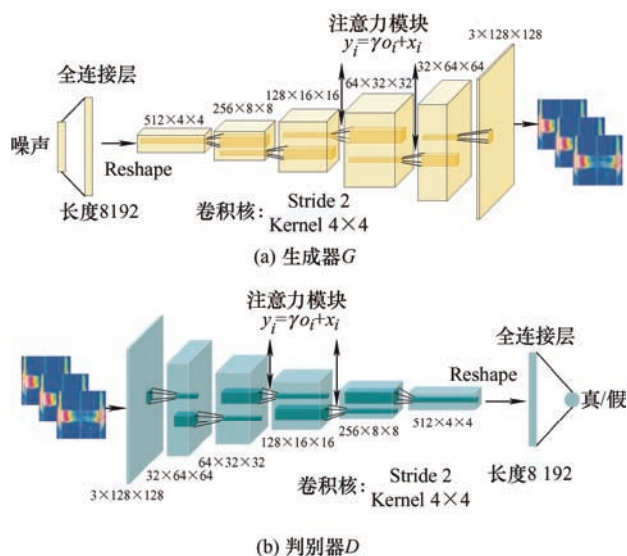


图 2 注意力嵌入的 GAN 模型结构

图 2 中关于注意力嵌入的 GAN 模型的具体构造细节如下。

(1) 生成器 G 构造。首先, 经全连接层对一维

高斯噪声进行扩展, 随后通过 reshape 操作转换为 $512 \times 4 \times 4$ 矩阵。最终, 通过对多个转置卷积层的上采样操作, 可将随机一维噪声输出为大小为 $3 \times 128 \times 128$ 的图像。在生成器中, 注意力模块被放置于第三和第四个卷积层之后, 以协调转置卷积所生成图像的不同位置的细节。此外, 使用高斯误差线性单元^[27](Gaussian error linear unit, GELU)作为每个卷积层的激活函数, 以进一步提供更为平滑的梯度更新, 其定义式如下

$$\text{GELU}(x) = 0.5x(1 + \tanh[\sqrt{2/\pi}(x + 0.044715x^3)]) \quad (8)$$

(2) 判别器 D 构造。判别器结构与生成器对称, 经过 5 层卷积特征提取后, 输入的 $3 \times 128 \times 128$ 的红外热成像图片被转换为 $256 \times 4 \times 4$ 特征张量。随后经 reshape 操作将其重构为一维数据, 最终输入全连接层中, 得到输出的值代表对于输入图像来源的判断结果。注意力模块被放置在与生成器相同的位置, 学习全局热力关联特征以提高生成图像的质量。在判别器中, 选择 Leaky 整流线性单元(Leaky rectified linear unit, LReLU)作为每层的激活函数。

2.3 双阈值训练机制开发

在传统 GAN 训练过程中, 通常需要判别器的性能在训练初期比生成器略微强大, 使得生成器能更好的学习到真实样本的特征。但性能过强的判别器对于生成样本的输出接近于 0, 其对生成器的梯度更新有着不利影响。此外, 通常需要设置冗余的训练迭代次数, 以确保生成样本的质量, 这导致了传统 GAN 训练的低效率和低适应性。

理想的对抗训练过程为判别器在训练初期能够快速更新参数, 而在中后期能够调整优化速度以避免阻碍生成器优化。此外, 在生成样本质量达到要求时, 训练过程能够自动停止。因此需要在不同训练阶段自适应采用训练策略, 以提高生成样本的质量和训练效率。开发了一种新的可自动采用不同训练策略的双阈值训练机制, 具体如下。

第一阶段, 结构相似度(Structural similarity, SSIM)的多次计算均值来源于生成器所输出一个 batch 数量的生成图片和相同数量真实图片之间的相互计算。当 SSIM 的多次计算均值小于第一级阈值时, 对抗训练采用双尺度固定学习率, 即判别器的学习率大于生成器, 以获取较为成熟的判别器参数。其中, SSIM^[28]用于评估转子-轴承系统实测红外热成像图片和生成图片之间的相似性, 其定义为

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \cdot \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \cdot \left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \right) \quad (9)$$

式中, $C_1 = (0.01V)^2$, $C_2 = (0.03V)^2$, $C_3 = 0.5C_2$; V 为像素值的范围; μ_x 、 μ_y 、 σ_x 、 σ_y 和 σ_{xy} 分别为真实样本 \mathbf{x} 和生成样本 \mathbf{y} 的统计学参数; SSIM 值越大, 说明图片之间的相似性越高。

第二阶段, 当 SSIM 的多次计算均值在第一级阈值和第二级阈值之间处于平稳状态时, 此时, 生成器学习率不变, 判别器自适应采用循环学习率来调节性能, 循环学习率定义如下所示

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \left(1 + \cos \left(\frac{T_{\text{cur}}}{T_{\max}} \pi \right) \right) \quad (10)$$

式中, η_t , η_{\max} 和 η_{\min} 分别为当前学习率、设定的最大和最小学习率; T_{cur} 为当次迭代, T_{\max} 为迭代循环周期。

第三阶段, 当 SSIM 的多次计算均值保持大于第二级阈值时, 说明此时生成样本质量符合要求, 训练自动停止。采用双阈值训练机制具有以下优点: ① 自适应地调整判别器的学习率以提高生成样本的质量; ② 训练过程能够自动停止以提高效率。

2.4 所提方法总体流程及主要步骤

所提方法总体流程如图 3 所示, 主要步骤如下。

步骤 1: 利用红外热敏相机采集转子-轴承系统在时变转速条件下不同健康状态的红外热成像图片, 并划分为小样本训练集和正常数量的测试集。

步骤 2: 构建双阈值注意力 GAN, 具体如下。

(1) 结合 Wasserstein 距离和梯度惩罚设计新型对抗损失函数, 避免训练过程中的梯度消失。

(2) 构建注意力嵌入的 GAN, 指导聚焦于红外热成像图片的全局热力关联特征。

(3) 开发双阈值训练机制, 以提高生成样本的质量和训练效率。

步骤 3: 针对小样本红外热成像图片, 利用双阈值注意力 GAN 生成新的辅助图片, 并从图片个体差异和整体分布等角度量化评估生成图片质量。

步骤 4: 利用生成的红外热成像图片辅助训练时变转速下的故障诊断模型, 通过对比验证所提方法的有效性和优越性。

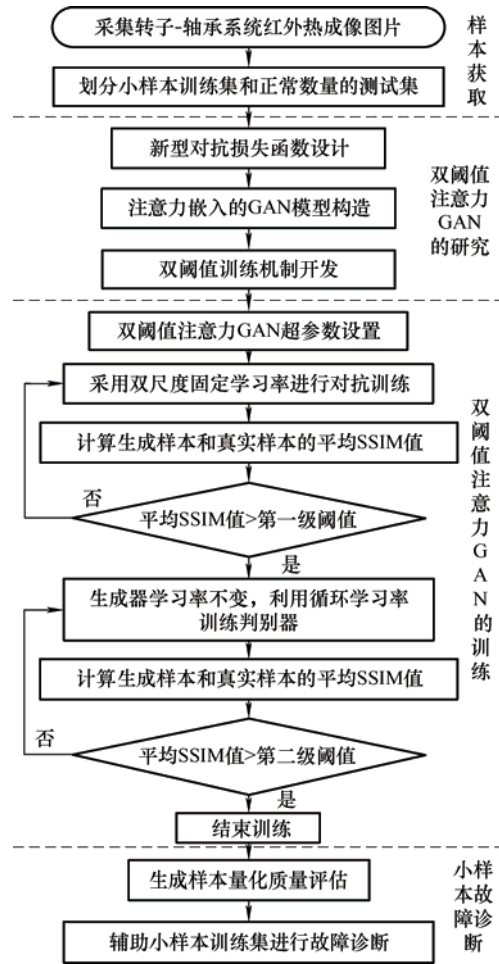


图 3 所提方法总体流程图

3 案例验证

3.1 红外热成像数据描述

由于缺乏工程现场的红外热成像数据, 所用的转子-轴承系统故障模拟试验台来源西北工业大学, 如图 4 所示, 其主要由红外热敏相机、数据采集装置、不平衡转子和测试轴承等组成^[29]。试验所用红外热成像图片由红外热敏相机(Hawk-1384, 分辨率 384×288 像素, 灵敏度为 0.05 °C)在环境温度为 19 °C 的时变转速情况下(900~3 000 r/min)连续采集。

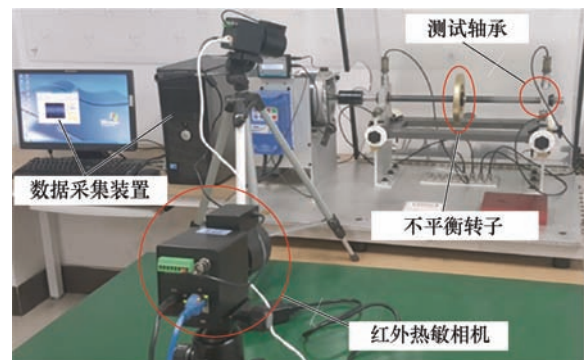


图 4 转子-轴承系统故障模拟试验台

转子-轴承系统的 8 种故障状态描述见表 1。在升速过程中连续拍摄得到每类故障 100 张红外热成像图片,为模拟小样本故障诊断和可能出现转速随机波动的工程情景,随机选取每类 10 张作为训练样本,而故障诊断的测试样本为每类 50 张。不平衡转子故障模拟和故障轴承分别如图 5a 和图 5b 所示。为了提高效率,人工选取感兴趣区域进行分析,如图 5c 所示。经图像变换处理,最终用于 GAN 训练和故障诊断的数据为边长 128 像素的正方形图片。

表 1 转子-轴承系统故障数据描述

不同故障状态	训练样本数量	测试样本数量
滚珠故障+平衡转子(C0)	10	50
滚珠故障+不平衡转子(C1)	10	50
复合故障+平衡转子(C2)	10	50
复合故障+不平衡转子(C3)	10	50
内圈故障+平衡转子(C4)	10	50
内圈故障+不平衡转子(C5)	10	50
外圈故障+平衡转子(C6)	10	50
外圈故障+不平衡转子(C7)	10	50

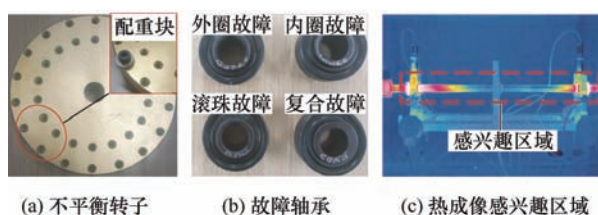


图 5 故障模拟部件以及红外热成像图片感兴趣区域

3.2 生成样本质量评估

为避免运行结果的偶然性和所选工况的随机性,重复 5 次的随机抽取表 1 中所述训练样本,用于训练双阈值注意力 GAN 以生成更多图片。案例运行环境及主要参数设置描述如下。CPU: i5-10400F@2.90 GHz; GPU: GTX1650; 运行内存: 16GB; 编程语言: Python3.6; 模型框架: Pytorch1.2; 双阈值训练机制中两级阈值分别为 0.60 和 0.92; 判别器和生成器学习率分别为 4×10^{-4} 和 2×10^{-4} ; 式(7)中梯度惩罚系数 $\lambda=1$; 式(10)中 η_{\max} 和 η_{\min} 分别为 4×10^{-4} 和 3.5×10^{-4} , T_{\max} 为 20。图 6 展示了某次运行时所提方法的生成图像与真实图像对比,可知生成的红外热成像图片与真实图片直观上较为相似。采用量化指标评估每次试验中所提方法(方法 1)生成样本的质量,其中随机选取生成样本每类 10 张与对应的真实训练样本(每类 10 张)进行量化指标的计算。同时,使用多种常见 GAN 方法进行质量

评估对比,包括使用新型对抗损失函数的自注意力 GAN(方法 2),使用新型对抗损失函数的深度卷积 GAN(方法 3),自注意力 GAN(方法 4),深度卷积 GAN(方法 5)和 ACGAN(方法 6),且所用对比 GAN 方法的网络层数,卷积核参数和激活函数层等设置基本一致。

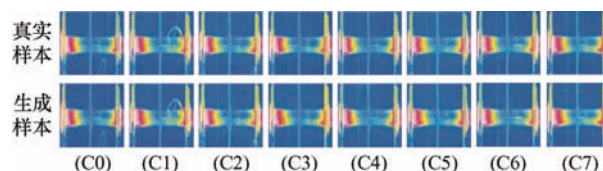


图 6 生成样本与真实样本的对比展示

首先,将 SSIM 应用于生成图片的个体差异评估,评估结果如图 7 所示。可知在 5 次重复试验中,所提方法生成的 8 种状态的平均 SSIM 值为 0.937 2,而其他 5 种方法的平均 SSIM 值分别为 0.922 6、0.904 5、0.917 1、0.879 3 和 0.866 3。因此,所提方法生成样本的质量在个体水平上优于对比方法。

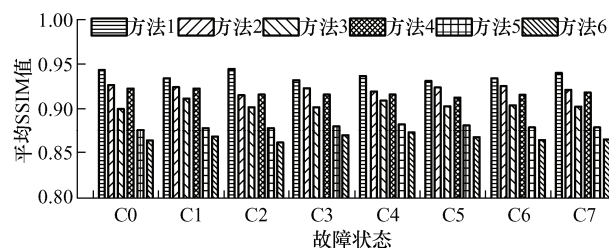


图 7 基于 SSIM 的个体差异评估结果

其次,将最大均值差异^[30](Maximum mean discrepancy, MMD)应用于生成图片的整体分布评估,MMD 值越小说明生成图片和真实图片的整体分布越接近,评估结果如图 8 所示。可知在 5 次重复试验中,所提方法生成的 8 种故障状态样本与真实样本之间的平均 MMD 值为 0.017 1,而对比方法的平均 MMD 值分别为 0.024 1、0.031 9、0.024 8、0.035 7 和 0.047 4。由此可知,相比其他方法,所提方法的 MMD 值更小,代表其生成样本在整体分布水平更接近真实样本。

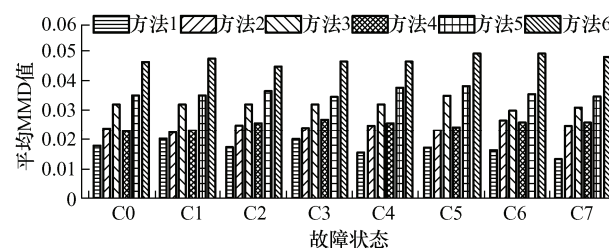


图 8 基于 MMD 的整体分布评估结果

最后,利用 t-分布领域嵌入算法对某次基于生成样本和真实样本混合提取的深度高维特征进行降

维可视化, 结果如图 9 所示。可知, 在其他三种未使用注意力模块的 GAN 中, 不同类别之间有少量样本混叠。因此, 所提方法的聚类效果更清晰, 表明嵌入注意力模块能够指导学习可分性更强的特征。

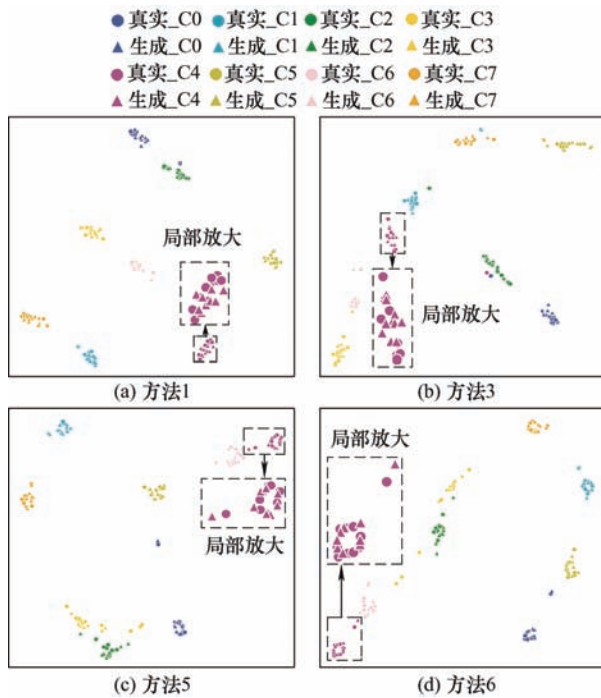


图 9 生成样本和真实样本的混合特征可视化结果

此外, 为了体现所开发双阈值训练机制的优越性, 在表 2 中记录了训练 8 类故障状态样本时不同的 GAN 方法所花费的平均迭代次数和平均训练时间。得出以下结论: ① 由方法 2 和方法 3、方法 4 和方法 5 两组对比可知, 在卷积网络中嵌入注意力模块会增加可训练参数, 导致训练时间延长; ② 由方法 1 和方法 2 对比可知, 所开发双阈值训练机制能提前自动停止训练以提高训练效率。图 10 中记录了使用 C0 条件样本训练不同 GAN 时 SSIM 值的变化趋势, 得出以下结论: ① 由方法 1 曲线趋势可知, 嵌入注意力模块和双阈值训练机制能够以更少的迭代次数生成更高质量图片; ② 由方法 2 和方法 4, 方法 3 和方法 5 两组对比得出, 新型对抗损失函数在训练初期能更快地找到优化目标。

表 2 不同 GAN 方法所需的平均迭代次数和平均训练时间

不同 GAN	平均迭代次数	平均训练时间/s
方法 1	1 492	434
方法 2	2 500	714
方法 3	2 500	503
方法 4	2 500	656
方法 5	2 500	514
方法 6	2 500	592

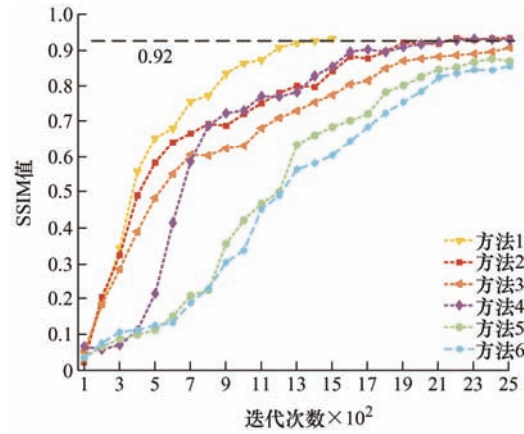


图 10 训练过程中的 SSIM 变化曲线

3.3 小样本故障诊断

利用生成的红外热成像图片辅助小样本的真实图片训练故障诊断模型, 以验证所提方法在小样本情景下故障诊断的有效性。故障诊断模型的分类器由 CNN 构建, 其中包括 4 个二维卷积层和最大池化层以及两个全连接层, 前两层卷积核大小为 5×5 ; 后两层卷积核大小为 3×3 ; 激活函数为 LReLU; 优化器为 Adam; 测试集为表 1 中所列。

为避免偶然性, 故障诊断试验重复进行 5 次, 诊断的平均精度和标准差如表 3 所示。可知, 在无 GAN 扩充的小样本情景(训练集 7)下, 仅使用小样本真实热成像图片的平均故障诊断准确率只有 72.75%, 经所提方法生成样本辅助训练后, 故障诊断的平均准确率提升至 96.15%, 且高于其他五种 GAN 方法的平均准确率, 诊断结果也更稳定。对比结果表明, 所提的双阈值注意力 GAN 能有效辅助小样本进行故障诊断, 有望克服故障样本匮乏问题。

表 3 基于不同训练集的故障诊断结果对比

不同训练集组成 (每类)	平均准确率(%)	标准差(%)
10 个真实样本+20 个方法 1 的生成样本	96.15	0.518 4
10 个真实样本+20 个方法 2 的生成样本	93.50	0.537 4
10 个真实样本+20 个方法 3 的生成样本	91.50	0.570 1
10 个真实样本+20 个方法 4 的生成样本	93.90	0.770 6
10 个真实样本+20 个方法 5 的生成样本	86.25	1.051 8
10 个真实样本+20 个方法 6 的生成样本	84.35	0.937 4
10 真实样本	72.75	2.910 1

此外, 为了进一步探索小样本数量对故障诊断性能的影响, 使用不同小样本数量的真实红外热成像图片来训练双阈值注意力 GAN。平均诊断结果如表 4 所示, 两个分类混淆矩阵如图 11 所示。结果表明, 当使用每类故障状态 8 个真实样本作为训练集

时,所提方法即能达到 90% 以上的平均诊断准确率。

表 4 不同数量真实样本训练所提方法的诊断结果对比

不同训练集组成(每类)	平均准确率(%)
6 个真实样本+24 个生成样本	82.75(1655/2000)
8 个真实样本+22 个生成样本	91.50(1830/2000)
10 个真实样本+20 个生成样本	96.15(1923/2000)
15 个真实样本+15 个生成样本	97.65(1953/2000)

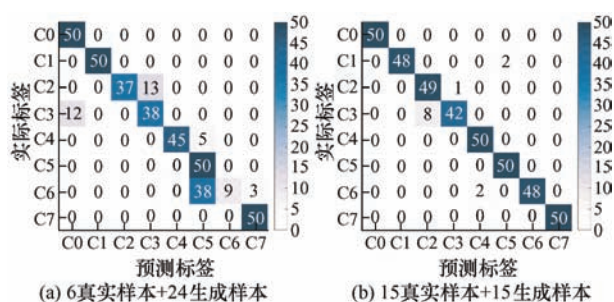


图 11 基于不同数量真实样本的诊断结果混淆矩阵

3.4 时变转速下红外热成像故障诊断的优势

开展时变转速下基于振动信号的故障诊断对比试验,以验证红外热成像的优越性。在图 4 所示试验台中安装加速度传感器,以获取时变转速下的振动信号,采样频率为 12 800 Hz。图 12 中展示了时变转速下采集的 8 种故障状态的振动信号,由波形可知振动信号具有较明显的非平稳特征。

利用时变转速下充足的振动信号训练一维 CNN 以进行对比故障诊断试验,一维 CNN 同样由 4 个一维卷积层和两个全连接层组成,其卷积核大小分别为 9×1 、 9×1 、 5×1 和 3×1 ,其他设置与 3.3 中的分类器完全相同。基于振动信号的训练集主要考虑两种形式:原始时域信号和经傅里叶变换提取的频域信号。采用滑动窗口采样法^[31]进行对一维信号进行样本划分,步长为 100,窗口大小设置为 2 304,总共获取 100 个长度为 2 304 的一维样本。对比试验的训练集设置和诊断结果如表 5 所示,测试样本为每类故障状态 50 个。由诊断精度可知,在时变转速条件下,案例中红外热成像图片受变转速工况影响较小,其故障诊断性能明显优于基于振动信号的分析。红外热成像的故障诊断方法具有较好的应用前景,但在实际工况中,该方法仍具有以下几个主要亟待提升之处:① 故障检测成本较高,受环境温度等其他热负荷影响较大;② 面对设备结构较为复杂的情况时,可能难以获取能够表征故障的图像;③ 热成像故障特征敏感性较差,难以检测突发故障。因此,需要更多先进的技术和现场数据已进一步提升红外热成像检测方法的工程应用价值。

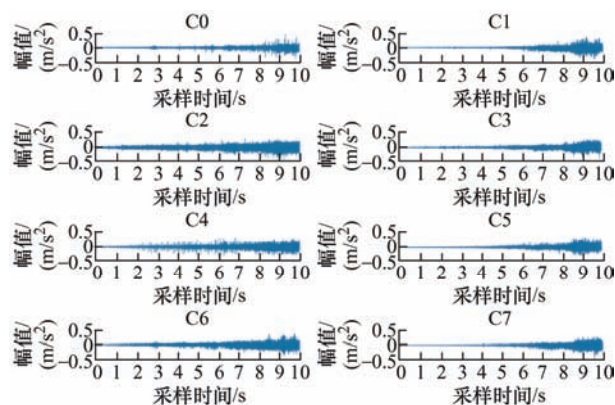


图 12 时变转速下 8 种故障状态的振动信号

表 5 不同数据形式的诊断结果对比

不同训练集形式	不同训练集组成(每类)	平均准确率(%)
红外热成像图片	10 个真实样本+20 个生成样本	96.15
红外热成像图片	50 个真实样本	99.80
原始时域信号	50 个真实样本	58.50
频域信号	50 个真实样本	71.65

4 结论

提出了一种双阈值注意力 GAN,结合红外热成像技术,解决了时变转速和小样本条件下的转子-轴承系统故障诊断难题。由对比试验表明,双阈值注意力 GAN 能生成高质量的红外热成像图片,在转子-轴承系统的小样本故障诊断方面比当前常用 GAN 方法具有更好性能。主要结论如下。

- (1) 基于 Wasserstein 和梯度惩罚设计的新型对抗损失函数,能够避免 GAN 训练中的梯度消失。
- (2) 通过嵌入注意力模块,可指导 GAN 提高对转子-轴承系统全局热力相关特征的学习能力。
- (3) 基于 SSIM 开发的双阈值训练机制,能有效提升生成图像的质量和训练效率。

参考文献

- [1] WEN Long, LI Xinyu, GAO Liang. A new reinforcement learning based learning rate scheduler for convolutional neural network in fault classification[J]. IEEE Transactions on Industrial Electronics, 2021, 68(12): 12890-12900.
- [2] 雷亚国, 许学方, 蔡潇, 等. 面向机械装备健康监测的数据质量保障方法研究[J]. 机械工程学报, 2021, 57(4): 1-9.
- LEI Yaguo, XU Xuefang, CAI Xiao, et al. Research on data quality assurance for health condition monitoring of

- machinery[J]. Journal of Mechanical Engineering, 2021, 57(4): 1-9.
- [3] 陈祝云, 钟琪, 黄如意, 等. 基于增强迁移卷积神经网络的机械智能故障诊断[J]. 机械工程学报, 2021, 57(21): 96-105.
- CHEN Zhuyun, ZHONG Qi, HUANG Ruyi, et al. Intelligent fault diagnosis for machinery based on enhanced transfer convolutional neural network[J]. Journal of Mechanical Engineering, 2021, 57(21): 96-105.
- [4] XING Saibo, LEI Yaguo, WANG Shuhui, et al. Distribution-invariant deep belief network for intelligent fault diagnosis of machines under new working conditions[J]. IEEE Transactions on Industrial Electronics, 2021, 68(3): 2617-2625.
- [5] 邵海东, 张笑阳, 程军圣, 等. 基于提升深度迁移自动编码器的轴承智能故障诊断[J]. 机械工程学报, 2020, 56(9): 84-90.
- SHAO Haidong, ZHANG Xiaoyang, CHENG Junsheng, et al. Intelligent fault diagnosis of bearing using enhanced deep transfer auto-encoder[J]. Journal of Mechanical Engineering, 2020, 56(9): 84-90.
- [6] GAO Yiping, GAO Liang, LI Xinyu, et al. A zero-shot learning method for fault diagnosis under unknown working loads[J]. Journal of Intelligent Manufacturing, 2020, 31(4): 899-909.
- [7] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Advances in Neural Information Processing Systems, 2014, 3: 2672-2680.
- [8] SHAO Siyu, WANG Pu, YAN Ruqiang, et al. Generative adversarial networks for data augmentation in machine fault diagnosis[J]. Computers in Industry, 2019, 40(9): 16-26.
- [9] DAI Jun, WANG Jun, HUANG Weiguo, et al. Machinery health monitoring based on unsupervised feature learning via generative adversarial networks[J]. IEEE/ASME Transactions on Mechatronics, 2020, 25(5): 2252-2263.
- [10] 肖雄, 肖宇雄, 张勇军, 等. 基于二维灰度图的数据增强方法在电机轴承故障诊断的应用研究[J]. 中国电机工程学报, 2021, 41(2): 738-749.
- XIAO Xiong, XIAO Yuxiong, ZHANG Yongjun, et al. Research on the application of the data augmentation method based on 2D gray pixel images in the fault diagnosis of motor bearing[J]. Proceedings of the CSEE, 2021, 41(2): 738-749.
- [11] WANG Rugen, ZHANG Shaohui, CHEN Zhuyun, et al. Enhanced generative adversarial network for extremely imbalanced fault diagnosis of rotating machine[J]. Measurement, 2021, 180: 109467.
- [12] LIU Shaowei, JIANG Hongkai, WU Zhenghong, et al. Data synthesis using deep feature enhanced generative adversarial networks for rolling bearing imbalanced fault diagnosis[J]. Mechanical Systems and Signal Processing, 2022, 163: 108139.
- [13] HE Zhiyi, SHAO Haidong, ZHONG Xiang, et al. An intelligent fault diagnosis method for rotor-bearing system using small labeled infrared thermal images and enhanced CNN transferred from CAE[J]. Advanced Engineering Informatics, 2020, 46: 101150.
- [14] LI Xin, SHAO Haidong, JIANG Hongkai, et al. Modified Gaussian convolutional deep belief network and infrared thermal imaging for intelligent fault diagnosis of rotor-bearing system under time-varying speeds[J]. Structural Health Monitoring, 2022, 21(2): 339-353.
- [15] 陈是抒, 彭志科, 周鹏. 信号分解及其在机械故障诊断中的应用研究综述[J]. 机械工程学报, 2020, 56(17): 91-107.
- CHEN Shiqian, PENG Zhike, ZHOU Peng. Review of signal decomposition theory and its applications in machine fault diagnosis[J]. Journal of Mechanical Engineering, 2020, 56(17): 91-107.
- [16] AMIN N, AMIN T G, MAHMOUD O, et al. Intelligent fault diagnosis of cooling radiator based on deep learning analysis of infrared thermal images[J]. Applied Thermal Engineering, 2019, 163: 114410.
- [17] JANSSENS O, VAN DE WALLE R, LOCCUFIER M, et al. Deep learning for infrared thermal image based machine health monitoring[J]. IEEE-ASME Transactions on Mechatronics, 2018, 23(1): 151-159.
- [18] JIA Zhen, LIU Zhenbao, VONG ChiMan, et al. A rotating machinery fault diagnosis method based on feature learning of thermal images[J]. IEEE Access, 2019, 7: 12348-12359.
- [19] LI Yongbo, DU Xiaoqiang, WAN Fangyi, et al. Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging[J]. Chinese Journal of Aeronautics, 2020, 33(2): 427-438.
- [20] CHOUDHARY A, MIAN T, FATIMA S. Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images[J]. Measurement, 2021, 176: 109196.
- [21] MAO Gang, ZHANG Zhongzheng, QIAO Bin, et al.

- Fusion domain-adaptation cnn driven by images and vibration signals for fault diagnosis of gearbox cross-working conditions[J]. Entropy, 2022, 24(1): 119.
- [22] ARJOVSKY M, BOTTOU L. Towards principled methods for training generative adversarial networks[EB/OL]. [2017-01-17], <https://arxiv.org/abs/1701.04862>.
- [23] ZHANG H, GOODFELLOW I, METAXAS D. Self-attention generative adversarial networks[EB/OL]. [2018-05-21], <https://arxiv.org/abs/1805.08318>.
- [24] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein GAN[EB/OL]. [2017-01-26], <https://arxiv.org/abs/1701.07875>.
- [25] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of Wasserstein GANs[EB/OL]. [2017-05-31], <https://arxiv.org/abs/1704.00028>.
- [26] SUN Guangcong, DING Shifei, SUN Tongfeng, et al. SA-CapsGAN: Using capsule networks with embedded self-attention for generative adversarial network[J]. Neurocomputing, 2021, 423: 399-406.
- [27] YANG Weijun, PANG Chengxin, HUANG Jinhai, et al. Sequence-to-point learning based on temporal convolutional networks for nonintrusive load monitoring[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 2512910.
- [28] LI Wei, ZHONG Xiang, SHAO Haidong, et al. Multi-mode data augmentation and fault diagnosis of rotating machinery using modified ACGAN designed with new framework[J]. Advanced Engineering Informatics, 2022, 52: 101552.
- [29] LI Yongbo, WANG Xianzhi, SI Shubin, et al. A new intelligent fault diagnosis method of rotating machinery under varying-speed conditions using infrared thermography[J]. Complexity, 2019, 2019(2): 2619252.
- [30] YANG Bin, LEI Yaguo, JIA Feng, et al. A polynomial kernel induced distance metric to improve deep transfer learning for fault diagnosis of machines[J]. IEEE Transactions on Industrial Electronics, 2020, 67(11): 9747-9757.
- [31] 孟宗, 关阳, 潘作舟, 等. 基于二次数据增强和深度卷积的滚动轴承故障诊断研究[J]. 机械工程学报, 2021, 57(23): 106-115.
- MENG Zong, GUAN Yang, PAN Zuozhou. Fault diagnosis of rolling bearing based on secondary data enhancement and deep convolutional network[J]. Journal of Mechanical Engineering, 2021, 57(23): 106-115.
-
- 作者简介: 邵海东(通信作者), 男, 1990 年出生, 博士, 副教授, 博士研究生导师。主要研究方向为故障诊断与寿命预测, 数据挖掘与信息融合。
E-mail: hdshao@hnu.edu.cn