

基于生成对抗网络的热泵系统故障诊断数据集扩充方法研究^①

孙哲^{②*} 金华强^{**} 顾江萍^{**} 黄跃进^{*} 王新雷^{***} 郑爱武^{****} 沈希^{③*}

(* 浙江工业大学机械工程学院 杭州 310023)

(** 浙江工业大学教育科学与技术学院 杭州 310023)

(*** 伊利诺伊大学香槟校区农业与生物工程学院 厄巴纳 IL61801)

(**** 加西贝拉压缩机有限公司 嘉兴 314011)

摘要 随着深度学习的发展,越来越多的热泵系统故障诊断方法引入深度学习技术并取得了较好的效果。基于深度学习的故障诊断技术需要依赖大量带有标记的故障数据,而现实中这类数据获取较为困难,这限制了智能诊断技术的应用。针对这一问题,本文提出利用生成对抗网络(GAN)学习故障数据的分布,并生成更多的标记数据,实现故障数据集的扩充。针对热泵系统运行数据结构复杂且不同故障间差异小给模型学习带来较大困难这一问题,本文提出利用热泵系统基准模型将运行数据转化为残差数据并作为训练数据,降低数据复杂度,增加差异性。利用 MMD 指标和 1-NN 指标对生成的数据进行分析,发现生成数据分布和真实数据接近,且利用残差数据训练的 GAN 模型质量更高。利用故障诊断方法对引入不同比例生成数据的模型训练结果进行分析,发现生成数据的引入可以提高数据量不足条件下的故障诊断精度。实验结果表明,基于 GAN 的数据扩充方法可有效降低智能诊断对标记数据的依赖,是一种应用前景广阔的技术。

关键词 生成对抗网络(GAN);深度学习;数据集扩充;热泵系统;故障诊断

0 引言

我国北方推行“煤改电”政策以来,热泵系统的使用量逐年增加。相关研究表明,热泵系统能源消耗约占建筑能耗一半以上^[1],而建筑能耗作为用能大户,在社会总能耗中占比超过 40%^[2]。热泵系统运行故障会导致其效率降低,造成大量能源浪费,及时、准确地诊断出相关故障可以最大限度地保证系统高效运行,达到节约能源的目的。

现阶段热泵系统故障诊断方法逐渐朝着智能化方向发展,众多研究提出了一系列基于数据驱动和深度学习的故障诊断模型,提高了诊断精度。文献[3]

构建了深度模型用于空气源热泵渐变故障诊断,较好实现了早期渐变故障的预警与诊断。文献[4]基于离散化贝叶斯网络诊断制冷机组,并在 ASHRAE RP-1043 数据集上验证,得到了较高的诊断精度。文献[5]利用循环神经网络构建诊断模型,实现空气源热泵热水器制冷剂泄漏亚健康的预警与诊断。文献[6]对近年来建筑能耗智能诊断方法做了相关综述,较为详细地描述了热泵系统智能故障诊断的研究现状。然而这些诊断方法都需要大量的、带有故障标记的训练数据实现模型训练。而在实际系统中,并没有那么多故障数据,这些数据也很难准确标记出故障类型,这一限制成为了智能诊断方法实际

① 国家自然科学基金(51076143),浙江省重点研发计划(2020C04010)和浙江省基础公益研究计划(LGG19E050020,LGG18E050024)资助项目。

② 男,1989年生,博士;研究方向:制冷/热泵系统智能故障诊断及节能优化技术;E-mail: sunzhe91@zjut.edu.cn。

③ 通信作者,E-mail: sx@zjut.edu.cn。

(收稿日期:2020-09-08)

应用的瓶颈之一。要想实现高精度的故障诊断,必须具有充足的、均衡的训练数据集,即每种故障都具有足够的数量。解决办法之一是利用有限的数量进行数据扩充,生成更多的高度仿真的数据,从而解决数据量不足、数据集不均衡的问题。因此,如何生成高质量的仿真数据成为了解决故障诊断实际应用的关键。

近年来一种新型智能算法——生成对抗网络(generative adversarial networks, GAN),在图像生成等领域得到了极大的关注,该算法利用自我博弈的方式高效地学习训练数据样本的概率密度分布,并利用学习到的参数生成更多的仿真样本。这一技术迅速成为生成模型的主力军,在数据生成质量上远优于其他数据生成算法,如自回归网络、变分自编码等。利用生成对抗网络有望实现热泵系统故障数据集的扩充,进而解决智能故障诊断对数据量的依赖问题。

一些学者在相关领域的扩充方向应用 GAN 取得了可喜的成果。文献[7]较早地在生物医学领域应用了 GAN,实现人脑核磁共振图像的生成,解决了生物医学图像数据短缺的问题。文献[8]同样利用 GAN 实现人脑核磁共振图像生成,以提高脑部肿瘤智能识别的精度。文献[9]将 GAN 引入语音识别领域,利用条件 GAN 生成更多的语音数据,解决带标记的语音数据量不足的问题。为提高生成数据质量,将语音数据转化到频域,以降低 GAN 的训练难度。文献[10]利用 GAN 实现图像数据集扩充,解决图像分类研究中因图像数据不足造成的模型表达能力差的问题。文献[11]将 GAN 引入机械故障诊断领域,利用深度卷积 GAN 实现机械运行数据集的扩充,解决机械故障诊断过程中因数据集不均衡导致诊断精度偏低的问题。

以上研究虽然证明 GAN 在数据集扩充方面具有一定优势,但尚无学者将这一技术应用到热泵系统故障数据集扩充方向。热泵系统故障数据相比于其他数据有着自身的特点。具体而言,不同故障数据间差异非常小,如健康系统的冷凝温度可能为 54.5℃,而冷凝器结垢系统的冷凝温度可能为 55.5℃。这种不明显的差异给不同类别的数据生成

带来了较大难度,其直接的结果是生成的各种类别数据可能相互混淆,无法达到提高诊断精度的目的。因此,本文针对热泵系统故障数据的这一特点,提出了利用热泵基准模型将运行数据转化为残差数据,并利用 GAN 学习残差数据,从而增大数据间的差异性。并针对 GAN 训练困难的问题,设计了具体可行的超参数方案和训练方法,实现了热泵系统故障数据集扩充。

1 基本理论

1.1 生成对抗网络

生成对抗网络(GAN)于 2014 年提出^[12],随后的几年内得到了领域内极大的关注。GAN 通过对抗的过程训练生成模型,包括 2 个神经网络,即判别网络和生成网络。训练的目标是使得生成网络 G 学习到真实数据的概率分布。

早期的 GAN 使用的全连接神经网络构建生成模型和判别模型。定义一组噪音变量 $p_z(z)$,并通过构建一个神经网络 $G(z; \theta_g)$ 将噪音变量映射到数据空间。然后,再构建一个神经网络 $D(x; \theta_d)$,其输出为单一的标量,表示 x 属于真实数据的概率。通过梯度下降的方式训练判别模型 D ,使其最大化正确判断数据真实与否的概率。同时,训练生成模型 G ,使其最小化 $\log(1 - D(G(z)))$ 。以上的训练过程表示,生成模型 G 和判别模型 D 在进行二元极大极小博弈,该博弈过程可以表示为

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

其中, p_{data} 是真实数据的分布, p_z 是隐空间的分布。

原始的 GAN 存在许多问题,包括训练困难、容易发散等。为提高 GAN 的使用性,后续研究逐渐提出了一系列新的版本,提高了训练稳定性,包括 Wasserstein GAN^[13] (WGAN) 以及它的改进版本 WGAN-GP^[14]。

WGAN 使用 Wasserstein 距离作为损失函数来训练 GAN,解决了当真实分布与生成分布不存在重叠时的梯度消失问题。Wasserstein 距离的定义如式

(2)所示。

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} \mathbb{E}_{(x, y) \sim \gamma} [\|x - y\|] \quad (2)$$

其中, $\Pi(P_r, P_g)$ 是所有以 P_r 和 P_g 为边界的联合分布集合。但是这一定义中的 $\inf_{\gamma \in \Pi(P_r, P_g)}$ 没有办法直接求解,因此将式(2)变换成式(3)的形式。

$$W(P_r, P_g) = \sup_{\|f\|_L \leq 1} \mathbb{E}_{x \sim P_r} [f(x)] - \mathbb{E}_{x \sim P_g} [f(x)] \quad (3)$$

这里函数 f 必须满足 Lipschitz 连续条件,即 $\|f\| \leq 1$ 。为了使其满足这一条件, WGAN 将判别网络 D 中的参数限制在 $[-c, c]$ 之间。

后续的研究发现 WGAN 有时同样会出现收敛失败的现象,导致这一现象的原因是强制限制了网络参数范围以满足 Lipschitz 连续条件^[14]。因此,该项研究提出一种利用梯度惩罚的策略来代替强制限制,并基于此策略提出了一种新型 GAN 变体——WGAN-GP。这一新型结构训练稳定、易于收敛,得到了广泛的应用。

1.2 生成数据质量评估

评价 GAN 模型是一项非常具有挑战性的工作。对于复杂的生成数据,不仅需要评价其与真实数据概率分布间的差异,也需要评价数据自身的多样性。目前尚没有公认的权威评价方法。文献[15]对常见的 GAN 评价方法做了详细的对比分析,得出核最大平均差异(kernel maximum mean discrepancy, MMD)和 1-最近邻(1-nearest neighbor, 1-NN)指标是 2 个非常优秀的评价指标,不仅可以用来区分真实对象和生成样本,还可以识别模式丢弃(modedropping)和模式坍塌(modecollapsing),以及检测过拟合。

Kernel MMD 指标可以定义为

$$\begin{aligned} MMD^2(P_r, P_g) &= \mathbb{E}_{\substack{x_r, x'_r \sim P_r \\ x_g, x'_g \sim P_g}} [k(x_r, x'_r) - 2k(x_r, x_g) + k(x_g, x'_g)] \end{aligned} \quad (4)$$

在一些固定的核函数 K 下, MMD 指标度量了真实分布 P_r 和生成分布 P_g 之间的差异,两分布差异越小,则 MMD 值越小。文献[16]提出的 parzen 窗估计是 MMD 的一种特殊形式。

1-NN 则用于成对的样本检验,用于评估两分布

的相似性。给定两组样本 $S_r \sim P_r^n$ 和 $S_g \sim P_g^n$, 满足 $|S_r| = |S_g|$, 可以计算在两分布上训练的 1-NN 分类器的 LOO (leave-one-out) 精度,其中 S_r 全部标记为正样本, S_g 全部标记为负样本。与其他精度度量不同的是,当 $|S_r| = |S_g|$ 非常大时, LOO 精度应服从约为 50% 的准确率,当两个分布相匹配时,即满足 50% 的准确率。当 LOO 低于 50%, 则说明生成分布过拟合真实采样分布。而当 LOO 远大于 50%, 甚至接近 100% 时,说明两分布差异很大。

本文将利用 MMD 和 1-NN 2 种评估指标对生成数据的质量进行评估,判断其与真实数据分布间的差异大小。

1.3 故障诊断验证方法

本文研究热泵系统故障数据集扩充方法,其目的是在标记数据不足的情况下提高诊断精度。因此,必须依靠一种诊断方法验证添加生成数据后诊断精度的变化。

本文选择前期提出的研究成果,一种基于深度学习的热泵系统故障诊断方法^[3]。该方法首先建立热泵系统基准模型,用于计算健康系统的理论状态值,如冷凝温度、排气温度等。然后利用实测状态值减去理论值,得到残差数据。最后利用卷积神经网络构建诊断模型,利用残差数据训练模型,并利用训练完成的模型实现热泵系统故障诊断。

本文构建两类训练数据集,一类数据集只包含真实数据,另一类数据集则是由真实数据和生成数据组成的混合数据集,具体数据集内的数据量详见后文第 3.5 节。

2 数据集扩充方法

2.1 学习数据选择

本文利用 GAN 学习数据的分布特性,并生成更多的仿真数据,用于扩充诊断算法的训练数据集。如第 1.3 节介绍,本文选用的诊断验证算法先要利用基准模型获得热泵系统的运行残差数据,后利用残差数据训练卷积神经网络,实现故障诊断。因此,数据集扩充方法既可以是扩充原始运行数据集,也可以扩充残差训练数据集。扩充原始数据集也需要

利用基准模型将其转化为残差数据供卷积神经网络训练。

热泵系统运行数据和残差数据有着较大的差别。运行数据是热泵系统的热物理参数,如压缩机排气压力、排气温度、冷凝器入口温度等。随着热泵系统运行工况的调整,这些热物理参数变化较大,一定程度上覆盖了因系统故障造成的参数变化,这些是热泵系统故障诊断困难的原因之一。直接利用 GAN 学习运行数据,就需要学习到不同故障数据间的差异,生成的数据也必须包含这些差异,否则无法达到提高诊断精度的目的。然而热泵系统结构复杂,具有强耦合、大滞后的特性,准确学习运行数据及故障间的差异具有较大难度。而残差数据则相对简单,所谓残差是指故障系统与健康系统在相同条件下各个参数的差值。在基准模型精度足够高的条件下,健康系统的残差数据应接近 0,不同故障系统的不同参数残差各不相同,使得故障数据间的差异性明显。利用 GAN 直接学习残差数据可以降低学习难度,提高数据生成质量。

但残差数据的生成需要依赖基准模型,使得生成残差的同时引入了基准模型自身的误差,模型精度越高,残差质量越好。本文选择文献[3]中的基准建模方法,这种基准模型融合卷积网络、编-解码器、循环网络,可以在热泵系统工况频繁迁移的条件下实现高精度建模,最大限度避免了残差计算过程中的误差引入。

2.2 算法构建

因 WGAN-GP 模型在训练稳定性方面的优势,本文以这一模型为基础,利用深度卷积网络搭建生成器和判别器,并根据数据集特性对模型以及参数设置进行优化改进,最终实现高质量的数据生成。

算法主要包括 2 个部分:(1) 生成网络 G , 采用上采样深度卷积网络;(2) 判别网络 D , 采用下采样深度卷积网络。

生成网络 G 的作用是接收输入的噪音数据,利用网络前向传播算法生成数据样本,数据结构为二维矩阵结构。生成网络 G 的训练目的是学习从噪音隐空间到样本数据 x 的映射关系。生成网络 G 的损失函数表示为式(5), p_g 表示生成数据的分布。

从损失函数可以看出,其计算值与判别器的诊断结果相关,这也是生成对抗两个过程不可分割的地方。

$$L_G = -\mathbb{E}_{x \sim p_g}[D(x)] \quad (5)$$

判别网络 D 的作用是接收真实数据样本和生成数据样本,利用网络前向传播算法最终给出数据是否属于真实数据的二元概率值,取概率较高的标签为最终判定值。判别网络 D 和生成网络 G 是在相互对抗的状态下完成训练的,判别网络 G 的损失函数表示为

$$L = \mathbb{E}_{\hat{x} \sim P_g}[D(\hat{x})] - \mathbb{E}_{x \sim P_r}[D(x)] + \lambda \mathbb{E}_{\hat{x} \sim P_{\omega}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (6)$$

其中, p_r 表示真实数据的分布, p_{ω} 表示真假数据间的抽样数据的分布。生成网络和判别网络在对抗学习中共同提高,最终使生成网络 G 可以生成概率分布非常接近真实数据分布的样本数据。

生成网络的目的是生成接近真实数据分布的高质量数据,用来迷惑判别网络。而判别网络则不断优化使得可以准确区分生成数据和真实数据。在这样一种对抗训练过程中,两个网络共同优化,最终使得生成网络可以生成非常接近真实数据的样本。为提高生成数据的质量,每个 GAN 模型只生成一种故障数据,多个 GAN 模型共同实现故障数据集的扩充。具体的 GAN 算法构架见图 1,运行数据先由基准模型转化为残差数据,而生成网络则直接由噪音数据生成残差数据。两种残差数据混合在一起输入判别网络,作为训练数据优化判别网络,使其可以更好地区分真实数据和生成数据。生成网络的训练则需要结合判别网络,将判别网络的参数固定使其不能优化,生成网络不断生成新的数据并将其标记成真实数据,利用判别网络诊断并计算损失值,利用反向传播算法优化生成网络。如此往复实现持续的对抗优化,直到训练轮数结束。

2.3 模型优化

GAN 模型训练困难,在现有模型基本结构的基础上需通过调节具体参数实现训练收敛,因此一些参数调整方面的优化对任务的实现具有重要作用。近年来,许多学者相继提出了一系列 GAN 的训练技巧,给 GAN 的实际应用带来了帮助,如 NIPS2016 以及 ICCV2017 上 Soumith Chintala 总结的 GAN 训练技

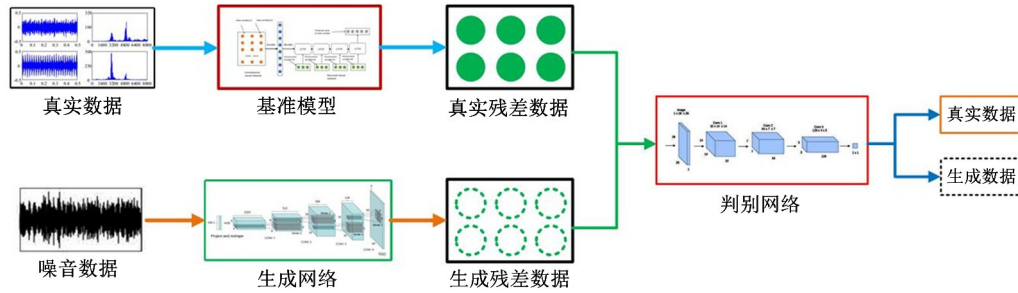


图 1 GAN 算法构架

巧。本文结合相关研究内容,根据自身研究特点,使用如下训练技巧。

(1) 除生成器第一层和判别器最后一层外,不使用全连接神经网络。

(2) 生成器激活函数选择方面,除最后一层外,均使用 rule 函数。最后一层根据数据的归一情况,若范围为 $[-1, 1]$ 则使用 tanh 函数,若为 $[0, 1]$ 则使用 sigmoid 函数。

(3) 判别器激活函数选择情况,除最后一层外,均使用 leakyrelu 激活函数。最后一层根据数据真伪标记法决定,若为传统 GAN 标记法,即 fake 数据标记为 1, valid 数据标记为 0, 则使用 sigmoid 函数;若使用带有梯度惩罚 (gradient penalty) 的标记法,即 fake 为 1, valid 为 -1 , dummy 为 0, 则使用 tanh 函数。

(4) 数据真伪的标记使用软标签,即不要简单地使用 1 或者 -1 来表示,而是添加 $[0, 0.1]$ 的随机噪音,将标记改为 $[-1, -0.9]$ 以及 $[0.9, 1]$ 之间的随机数。

(5) 在生成每一层卷积层之后,添加一层 batch-normalization (BN) 层。

(6) 若使用 WAGN-GP 模型,在判别器中不要使用 batchnormalization 层。

(7) 无论是生成器还是判别器,卷积核的数据应比较多,本文实验中,使用的最大卷积核数量达到了 256。

(8) 无论是生成器还是判别器,均不使用池化层。

(9) 判别器不使用太深的卷积网络,通常 2 ~ 3 层的深度最好,太深的判别器范围诊断效果不好。

(10) 每训练一轮生成器,通常需要训练 n 次判

别器,本文设置 $n = 3$ 。

本文所用 GAN 模型的生成器和判别器均使用深度卷积网络,具体结构参数见表 1 和表 2。利用以上训练技巧和优化的参数,可以实现热泵系统残差数据的高质量生成。

表 1 生成器网络结构

网络层	核大小	步长	输出大小
全连接层	-	-	(None, 3200)
Reshape 层	-	-	(None, 5, 5, 128)
反卷积层	4×4	$[2, 2]$	(None, 10, 10, 128)
BN 层	-	-	(None, 10, 10, 128)
反卷积层	4×4	$[2, 1]$	(None, 20, 10, 64)
BN 层	-	-	(None, 20, 10, 64)
反卷积层	4×4	$[1, 1]$	(None, 20, 10, 32)
BN 层	-	-	(None, 20, 10, 32)
反卷积层	4×4	$[1, 1]$	(None, 20, 10, 1)

表 2 判别器网络结构

网络层	核大小	步长	输出大小
卷积层	4×4	$[2, 2]$	(None, 10, 5, 64)
Drop 层	-	-	(None, 10, 5, 64)
卷积层	4×4	$[2, 1]$	(None, 5, 5, 128)
Drop 层	-	-	(None, 5, 5, 128)
Flatten 层	-	-	(None, 3200)
全连接层	-	-	(None, 1)

3 实验与评估

3.1 数据集来源

本文实验数据来自项目组搭建的热泵系统故障诊断实验平台,平台可以分成 3 个子系统,即热泵子系统、电气子系统和软件子系统,其结构原理见图 2。

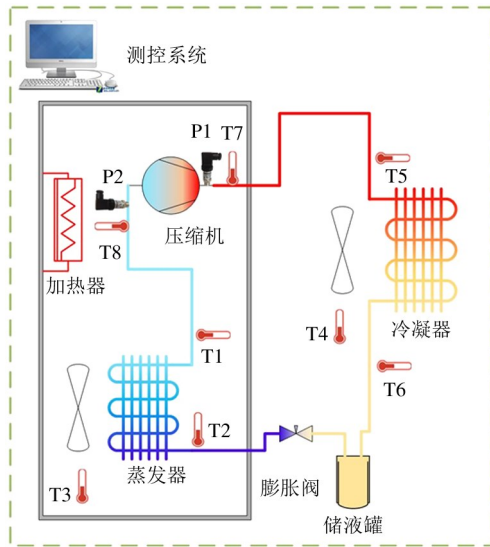


图 2 实验平台结构图

热泵子系统的详细介绍见表 3,冷媒选择 R134A。电气子系统包括数据采集、机构控制和 DA 输出。数据采集系统采用研华数据采集系统。软件子系统采用 LabVIEW2018 编程。每间隔 30 s 进行一次运行数据采集,采集完毕的数据存储在 MySQL 数据库。

表 3 热泵子系统介绍

组件	描述
压缩机	定频往复压缩机 排气量:12 cm ³
冷凝器 (室内单元)	管式翅片换热器 管路行/列: 2/10 宽/长/高(mm): 250/120/240 换热面积(m ²): 3.4
蒸发器 (室外单元)	管式翅片换热器 管路行/列: 1/9 宽/长/高(mm): 250/120/240 换热面积(m ²): 2
膨胀阀	热力膨胀阀(利用外接氮气控制)

实验平台可实现 8 种系统状态模拟,包括健康状态、冷凝器结垢、蒸发器结垢、制冷剂泄漏、冷凝器蒸发器同时结垢、冷凝器结垢且制冷剂泄漏、蒸发器结垢且制冷剂泄漏和冷凝器蒸发器结垢且制冷剂泄漏。采集的数据参数见表 4。

表 4 采集数据变量

数据子集	变量
环境变量	室外温度,室内温度
控制变量	压缩机转速,膨胀阀开度
状态变量	排气温度,吸气温度,冷凝温度,蒸发温度,冷凝器出口温度,蒸发器入口温度,排气压力,吸气压力

3.2 模型训练

GAN 训练需要利用一定数量的真实数据,实验中对每类故障单独训练。每类故障选择 500 组样本构成训练集, batchsize 设置为 200,共训练 10 万轮。每完成 1000 轮,对生成模型进行保存,并利用生成模型生成一批数据,用于 GAN 模型质量的评估。

本文提出利用 GAN 直接学习残差数据代替学习原始运行数据,以此用来降低 GAN 训练难度。为验证这一方法的有效性,在 GAN 训练中分别让其学习残差数据和原始数据,并对生成的数据质量进行评估。

算法利用 Python 3.6 编程,开发环境使用 Pycharm 2017 版本。深度学习框架使用 Keras 2.2.4,这是一种成熟的深度学习顶层封装框架,其底层使用 TensorFlow 运算,最新版的 TensorFlow 框架已经融合 Keras。

硬件环境为一台图形工作站,CPU 为 Intel Xeon E5-1650 v3,GPU 为 NVIDIA Quadro K2200,工作站内存为 16 GB,操作系统为 Windows 7 X64,训练算法运行在 GPU 上。

3.3 基于 MMD 指标的评估验证

为验证 GAN 生成数据的质量,本节利用 MMD 指标对生成数据与真实数据间分布差异进行评估。本文提出利用 GAN 模型学习热泵系统残差数据代替原始数据,用于提高数据生成质量。因此,本节利用 MMD 分别评估两种 GAN 模型生成数据(原始生成数据和残差生成数据)的质量,用于验证利用残差数据训练 GAN 的有效性。MMD 指标是考量两组数据分布差异的一种方法,MMD 值越小,两组分布差异越小,可以用来验证生成数据和真实数据的接近程度。

本节对实验不同训练轮数对应的生成数据进行

MMD 分析,对比原始数据生成实验和残差数据生成实验在相同训练轮数时生成数据的质量,实验结果见图 3。图中展示了两组 GAN 训练实验中各类故障生成数据随训练轮数增加 MMD 评分的变化。其中,health 表示健康系统数据、co 表示冷凝器结垢故障、ev 表示蒸发器结垢故障、re 表示制冷剂泄漏故障、c_e 表示蒸发器结垢冷凝器结垢故障、r_c 表示冷凝器结垢制冷剂泄漏故障、r_e 表示蒸发器结垢制冷剂泄漏故障、r_c_e 表示蒸发器结垢冷凝器结垢制冷剂泄漏故障。

两组实验均进行 10 万轮的对抗训练。由图 3 可知,随着训练轮数的增加,所有故障生成数据的 MMD 值均呈现下降趋势,说明对抗生成训练持续优化,证明前文所描述的 GAN 网络结构及其超参数设定是可行的。对比不同故障数据,原始数据生成实验中的健康生成数据的 MMD 值相比于其他数据偏低,但仍然呈现下降趋势,在训练轮数接近 10 万轮时,MMD 值约为 0.32,属于可接受范围,其他故障生成数据的质量则更高。对比原始数据生成实验和残差数据生成实验,发现残差数据生成实验的 MMD 值普遍低于原始数据生成实验,数据质量更加接近真实数据,造成这一现象的主要原因在于数据自身的差异性。原始数据不同故障以及不同时刻下数据自身差异性较小,例如,冷凝器结垢亚健康和健康系统对比,冷凝温度变化幅度小于 2℃,而冷凝温度的基数约为 55℃,波动比例约为 4%。而相对于残差数据,冷凝器结垢故障和健康系统对比,前者的冷凝温度残差约为 -0.2℃,而后者的冷凝温度残差在 [-0.01, 0.01] 之间,波动比例超过 20 倍。这一现

象说明不同时刻以及不同健康程度下,残差数据的差异性更加明显,使得 GAN 生成数据更加容易。

从图 3 可以得出结论,残差数据生成实验的数据相比于原始数据生成实验质量更优。接下来对 10 万轮训练完毕的数据进行交叉验证,即分析不同故障数据间的 MMD 值,用于验证生成数据对不同故障标签的区分程度。

对于真实数据而言,同一故障类别的数据分布相同,MMD 值较低;而不同故障类别的数据分布不同,MMD 值较高。系统生成数据也符合这一情况。表 5 展示了原始数据生成实验的 MMD 指标交叉验证数据。从表中可以看出,相同类别的生成数据与真实数据之间的 MMD 值明显低于不同类别的 MMD 值。证明生成的数据分布跟对应类别的真实数据分布接近,跟不同类别的真实数据分布差异较大。

同样地,表 6 展示了残差数据生成实验的 MMD 交叉验证数据。与原始数据生成实验相同,残差数据生成实验中,相同类别的生成数据和真实数据 MMD 值较低,不同类别的生成数据和真实数据 MMD 值较高。且同一类别的 MMD 值普遍比对应的原始数据生成实验的 MMD 值低,再次证明了残差数据生成实验的优势。

通过利用 MMD 指标对两组实验进行验证分析发现,无论是不同训练轮数下的 MMD 值还是交叉验证的 MMD 值,残差数据生成实验的生成质量均优于原始数据生成实验。因此,从 MMD 评分的角度来看,利用残差数据生成是一种更优的方法。

表 5 原始数据生成实验 MMD 交叉验证表

	health	co	ev	re	c_e	r_c	r_e	r_c_e
health	0.34	0.75	0.70	0.75	0.78	0.75	0.69	0.71
co	0.86	0.17	0.80	0.88	0.52	0.56	0.75	0.58
ev	0.81	0.86	0.10	0.84	0.64	0.85	0.59	0.57
re	0.85	0.88	0.77	0.35	0.79	0.66	0.38	0.62
ce	0.84	0.65	0.63	0.87	0.09	0.75	0.72	0.45
rc	0.83	0.58	0.79	0.72	0.68	0.19	0.63	0.53
re	0.81	0.82	0.55	0.59	0.68	0.69	0.14	0.48
rce	0.76	0.63	0.53	0.74	0.44	0.55	0.48	0.14

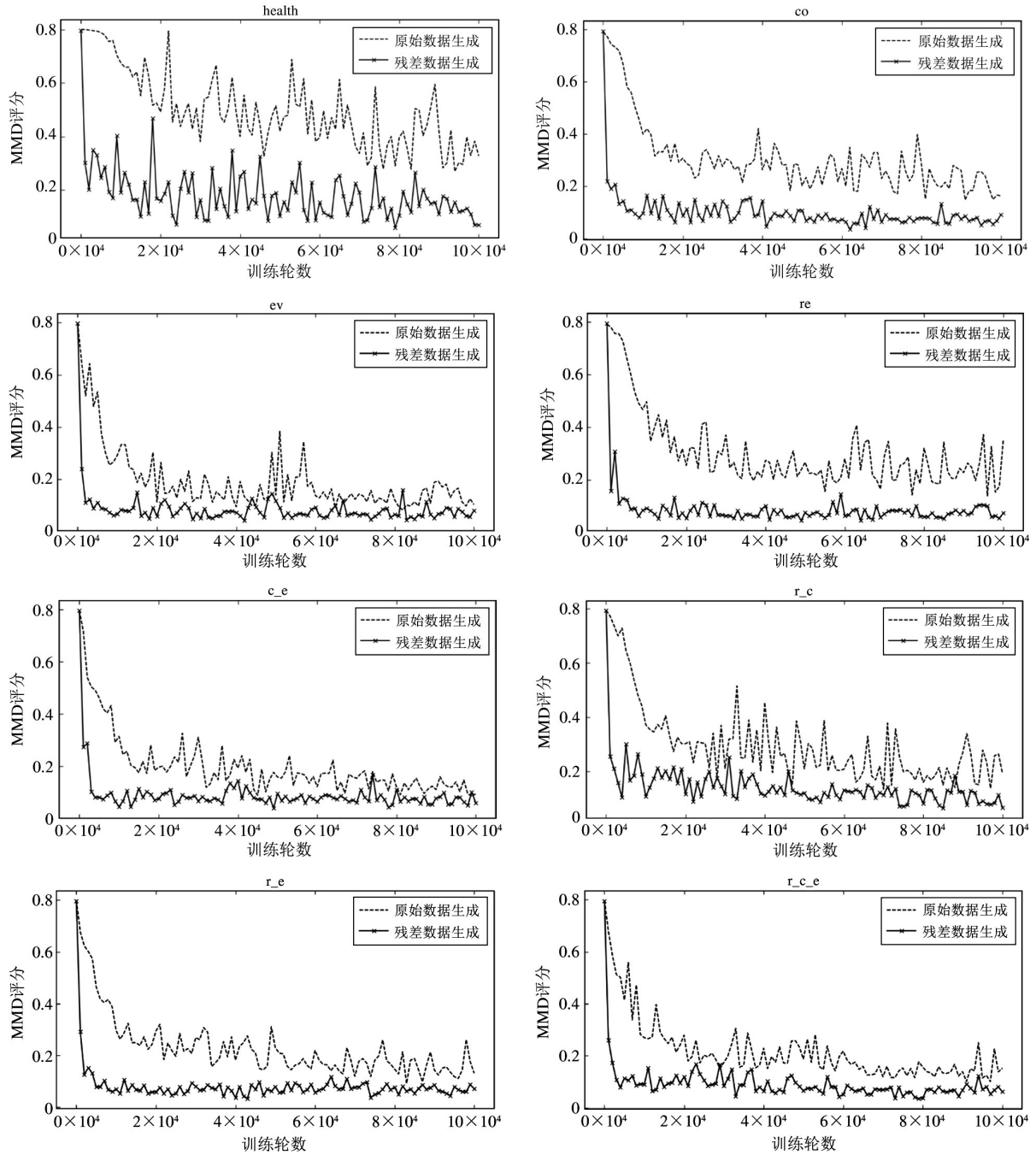


图3 基于 MMD 指标的评估结果对比图

表6 残差数据生成实验 MMD 交叉验证表

	health	co	ev	re	c_e	r_c	r_e	r_c_e
health	0.06	0.96	0.83	0.81	0.83	0.80	0.75	0.70
co	0.96	0.08	0.98	1.00	0.52	0.62	0.94	0.67
ev	0.86	0.95	0.09	0.85	0.76	0.91	0.61	0.61
re	0.86	0.99	0.83	0.07	0.90	0.80	0.45	0.73
ce	0.87	0.55	0.78	0.92	0.06	0.72	0.84	0.49
rc	0.86	0.64	0.90	0.81	0.68	0.07	0.77	0.53
re	0.83	0.94	0.64	0.50	0.83	0.81	0.07	0.62
rce	0.81	0.71	0.62	0.80	0.48	0.58	0.63	0.05

3.4 基于 1-NN 指标的评估验证

GAN 模型质量评估尚处于开放研究状态,没有一种权威的指标可以全面评价生成数据的质量。为更加准确地评估实验结果,本文利用 MMD 指标的同时使用 1-NN 指标再次评估,并综合评估结果进行分析,以期得出更加真实的结论。1-NN 指标与 MMD 不同,前者的理想值是 0.5,在 $[0.5, 1]$ 之间时,值越低越好,但是低于 0.5 则说明生成数据存在

过拟合。

利用评估指标对不同训练轮数的 GAN 生成数据质量进行评估,并将评估结果列于图 4。从图中可以看出,原始数据生成实验的结果并不理想,随着训练轮数的增加,部分故障数据的 1-NN 值存在上升情况,而其他故障数据也没有表现出良好的单调性。在训练 2 万轮之后,数据的 1-NN 值基本稳定,对抗训练不再能够继续优化。而残差数据训练实

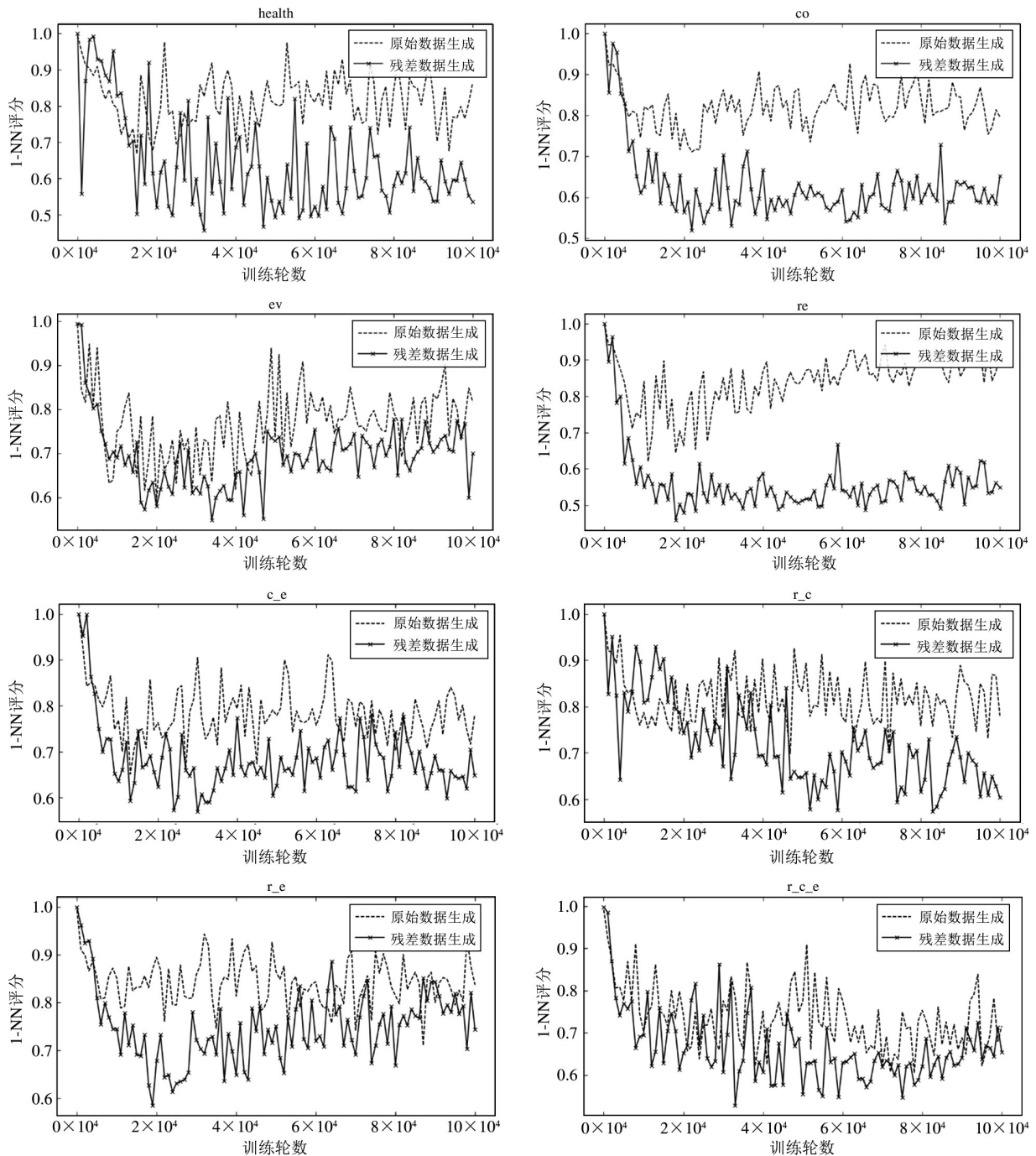


图 4 基于 1-NN 指标的评估结果对比图

验中,虽然下降幅度也不太明显,但整体 1-NN 值较原始数据生成实验的更低,更加接近 0.5 的理想状态。说明残差数据生成实验能更好地拟合真实数据分布,数据质量更高。产生这一现象的原因是残差数据自身更加明显的差异性导致的,这一差异性降低了模型的学习难度。

对 10 万轮训练完毕的数据进行交叉验证,分析不同故障数据间的 1-NN 值,验证不同故障标签生

成数据的区分程度。表 7 展示了原始数据生成实验的 1-NN 指标交叉验证数据。从表中可以看出,相同类别的生成数据与真实数据之间的 1-NN 值明显低于不同类别的 1-NN 值。证明生成的数据分布跟对应类别的真实数据分布接近,而跟不同类别的真实数据分布差异较大。同时发现,同类别的数据的 1-NN 值均大于 0.75,证明生成数据质量不是很理想。

表 7 原始数据生成实验 1-NN 交叉验证表

	health	co	ev	re	c_e	r_c	r_e	r_c_e
health	0.87	1.00	0.99	1.00	1.00	1.00	1.00	1.00
co	1.00	0.82	1.00	1.00	0.99	1.00	1.00	1.00
ev	1.00	1.00	0.82	1.00	1.00	1.00	1.00	1.00
re	1.00	1.00	1.00	0.94	1.00	1.00	1.00	1.00
ce	1.00	0.99	1.00	1.00	0.78	1.00	1.00	1.00
rc	1.00	1.00	1.00	1.00	1.00	0.78	1.00	0.99
re	1.00	1.00	1.00	0.97	1.00	1.00	0.83	1.00
rce	1.00	1.00	1.00	1.00	1.00	0.96	1.00	0.70

表 8 展示了残差数据生成实验的 1-NN 指标交叉验证数据。从表中看出,同样是相同类别的生成数据的 1-NN 值较低,不同类别的 1-NN 值较高,证明 GAN 模型可以针对不同类型数据进行区分生成。观察同类别的 1-NN 值发现平均在 [0.58, 0.7] 之

间,优于原始数据生成实验,再次证明残差数据生成方式的优势。

经过 1-NN 指标的验证,可以得出同样的结论。利用残差数据生成方法实现数据集扩充,得到的数据质量更高,更加接近真实分布。

表 8 残差数据生成实验 1-NN 交叉验证表

	health	co	ev	re	c_e	r_c	r_e	r_c_e
health	0.58	1.00	1.00	1.00	1.00	1.00	1.00	1.00
co	1.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00
ev	1.00	1.00	0.71	1.00	1.00	1.00	1.00	1.00
re	1.00	1.00	1.00	0.58	1.00	1.00	0.96	1.00
ce	1.00	1.00	1.00	1.00	0.67	1.00	1.00	1.00
rc	1.00	1.00	1.00	1.00	1.00	0.65	1.00	0.98
re	1.00	1.00	1.00	0.99	1.00	1.00	0.70	1.00
rce	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.65

3.5 基于故障诊断精度的评估验证

第 3.3 节和第 3.4 节中分别利用 MMD 指标和 1-NN 指标对生成数据与真实数据分布间的差异进行了评估,证明利用 GAN 生成的数据与真实数据分布相近,且残差生成实验的质量更高。本文研究数

据集扩充的目的是提高热泵系统故障诊断精度,因此有必要进一步利用诊断方法验证添加生成数据对最终诊断精度的影响。

本节共设计两组实验从两个维度验证生成数据添加对诊断精度的影响。每组实验中又分为利用原

始数据生成实验数据和残差数据生成实验数据进行验证的两类。

实验 1 中训练数据集包含 8 类故障数据子集, 每类故障数据子集均包含 500 组真实数据样本, 以及一定数量的生成数据样本。按每类生成数据样本的添加量不同, 分别构成 5 个混合训练数据集, 包含的每类生成数据样本分别为 0 组、500 组、1000 组、1500 组、2000 组。训练的参数设置及训练轮数参考文献[3]中设置。训练完成的模型利用未参与训练的测试数据集测试诊断精度, 测试数据集包含 8 类真实故障数据, 每类 2000 组样本, 实验结果见图 5。可以看出, 随着生成数据数量的增加, 整体诊断准确率呈现先增加后减少。在真实数据和生成数据比例为 1:1 时, 整体诊断准确率最高。同时发现, 在训练数据量较低时, 即便全部是真实数据, 诊断准确率也非常低。对比原始数据生成实验和残差数据生成实验发现, 利用生成的残差数据扩充训练数据, 其诊断准确率的提高效果远高于原始数据扩充的效果。随着生成数据的不断增加, 整体准确率会出现少量下降, 说明数据生成质量依然存在干扰, 不断增加生成数据的同时引入了干扰项。

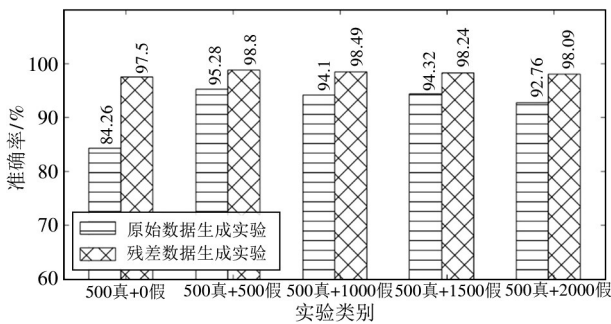


图 5 固定真实数据量的诊断验证实验

实验 2 的设计与实验 1 类似, 不同之处在于混合训练数据集中每类故障数据量固定为 2000 组, 其中真实数据与生成数据按 4:0、3:1、1:1、1:3、0:4 的 5 种比例组合, 构成 5 个不同的训练数据集。同样利用每类 2000 组未参与训练的真实数据进行验证, 结果见图 6。可以看出, 原始数据生成实验中, 随着真实数据比例的降低, 整体诊断精度下降, 说明利用生成数据代替真实数据会在一定程度上牺牲准确率。而残差数据生成实验中, 随着真实

数据比例的降低, 整体诊断精度下降并不明显, 仅在全部使用生成数据训练时存在明显的准确率下降。

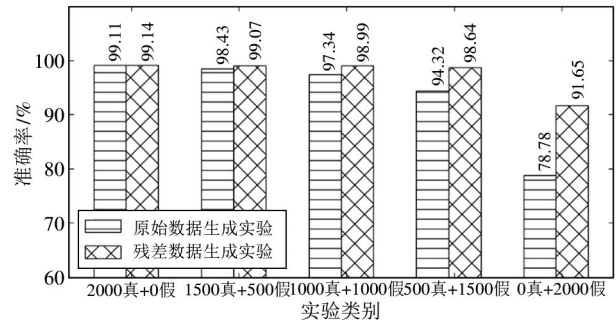


图 6 固定总体数据量的诊断验证实验

以上两组实验说明, 利用 GAN 进行数据集扩充是具有一定意义的。在数据量有限的情况下, 利用生成数据增加训练数据量可以达到提高诊断精度的目的。而对比原始数据生成实验和残差数据生成实验发现, 残差数据生成实验的效果明显优于原始数据生成实验, 这一结论与第 3.3 节和第 3.4 节的结论是一致的。

4 结论

随着近年来深度学习的迅速发展, 越来越多的热泵故障诊断方法引入深度学习技术, 得到了很好的效果。但基于深度学习的故障诊断依赖大量带有标记的故障数据, 而实际上热泵系统的故障数据非常难获取, 给深度学习的应用带来困难。本文提出利用新兴的智能算法——生成对抗网络 (GAN), 实现热泵系统故障数据集的扩充, 解决训练数据不足、数据集不均衡等问题。文章的主要工作及贡献总结如下。

(1) 针对热泵系统运行数据间差异较小且复杂度高问题, 提出利用热泵系统基准模型将运行数据转化成残差数据, 增大数据间差异化, 降低数据复杂度。利用 GAN 学习残差数据的分布, 将生成的残差数据用于故障分类器的训练, 实现故障诊断精度的提高。

(2) 针对 GAN 模型训练困难、易出现模式坍塌等问题, 提出一系列训练技巧来提高模型训练稳定性。并针对热泵数据自身特点, 设计并优化的 GAN

模型中生成器和判别器的结构参数,真正实现稳定、高效的故障数据集扩充。

(3)为验证生成数据和真实数据间的分布差异,利用 MMD 和 1-NN 指标对生成数据进行验证。结果显示,随着 GAN 训练轮数的增加,MMD 和 1-NN 指标均呈现下降趋势,证明训练过程处于持续优化。利用交叉验证,同类故障数据间的 MMD 和 1-NN 指标均远小于不同类别的指标值,证明 GAN 可以准确区分故障类别。同时,利用残差数据训练的 GAN 模型在两类指标上均优于原始数据训练的模型,证明利用残差数据训练 GAN 是一种有效的优化方法。

(4)利用诊断方法对引入生成数据的诊断实验进行验证。在含有固定 500 组真实样本的验证实验中,随着生成数据数量的增加,诊断精度先提高后轻微降低,在真实数据和生成数据量为 1:1 时,诊断精度最高。在固定 2000 组样本总量的验证实验中,随着生成数据比例的增加,诊断精度轻微下降,但并不明显,仅在全使用生成数据时精度明显下降。证明基于 GAN 的数据扩充对诊断精度的提高具有积极意义。

综上所述,本文提出的热泵系统故障数据集扩充方法对数据不足情况下的故障诊断精度具有明显的优化效果,可一定程度上降低对标记数据的依赖,扩大智能诊断方法的应用范围,是一种应用前景广阔的方法。而作为数据生成质量的重要保证,生成模型评价方法至关重要,目前尚无成熟评价体系,因此可作为未来研究方向之一,尝试构建从距离分布、多样性以及热泵数据独有特性保留程度等多方面综合评价的体系,实现更为全面的生成模型质量评估。

参考文献

- [1] Rashid S A, Haider Z, Chapal H S M, et al. Retrofitting low-cost heating ventilation and air-conditioning systems for energy management in buildings[J]. *Applied Energy*, 2019, 236: 648-661
- [2] Wei Y, Zhang X, Shi Y, et al. A review of data-driven approaches for prediction and classification of building energy consumption[J]. *Renewable and Sustainable Energy Reviews*, 2018, 82: 1027-1047
- [3] Sun Z, Jin H, Gu J, et al. Gradual fault early stage diagnosis for air source heat pump system using deep learning techniques [J]. *International Journal of Refrigeration*, 2019, 107: 63-72
- [4] Wang Y, Wang Z, He S, et al. A practical chiller fault diagnosis method based on discrete Bayesian network[J]. *International Journal of Refrigeration*, 2019, 102: 159-167
- [5] Sun Z, Jin H, Gu J, et al. Studies on the online intelligent diagnosis method of undercharging sub-health air source heat pump water heater[J]. *Applied Thermal Engineering*, 2020, 169: 114957
- [6] Zhao Y, Li T, Zhang X, et al. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: advantages, challenges and the future [J]. *Renewable and Sustainable Energy Reviews*, 2019, 109: 85-101
- [7] Calimeri F, Marzullo A, Stamile C, et al. Biomedical data augmentation using generative adversarial neural networks [C] // *Artificial Neural Networks and Machine Learning*, Alghero, Italy, 2017: 626-634
- [8] Bowles C, Chen L, Guerrero R, et al. GAN augmentation: augmenting training data using generative adversarial networks[EB/OL]. <http://arXiv.org/abs/1810.10863>, (2018-10-25), [2020-09-08]
- [9] Qian Y, Hu H, Tan T. Data augmentation using generative adversarial networks for robust speech recognition [J]. *Speech Communication*, 2019, 114: 1-9
- [10] 王海文, 邱晓晖. 一种基于生成式对抗网络的图像数据扩充方法[J]. *计算机技术与发展*, 2020, 30(3): 1-9
- [11] Du Y, Zhang W, Wang J, et al. DCGAN based data generation for process monitoring [C] // *2019 IEEE 8th Data Driven Control and Learning Systems Conference*, Dali, China, 2019: 410-415
- [12] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C] // *Neural Information Processing Systems*, Montreal, Canada, 2014: 2672-2680
- [13] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN [EB/OL]. <http://arXiv.org/abs/1701.07875>, (2017-01-26), [2020-09-08]
- [14] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs [C] // *Advances in Neural Information Processing Systems*, Long Beach, USA, 2017:

5767-5777

- [15] Xu Q, Huang G, Yuan Y, et al. An empirical study on evaluation metrics of generative adversarial networks[EB/OL]. <http://arXiv.org/abs/1806.07755>, (2018-01-19), [2020-09-08]

- [16] Gretton A, Borgwardt K, Rasch M J, et al. A kernel method for the two-sample problem[EB/OL]. <http://arXiv.org/abs/0805.2368>, (2008-05-15), [2020-09-08]

Data augmentation for heat pump systems fault diagnosis based on generative adversarial networks

Sun Zhe^{*}, Jin Huaqiang^{**}, Gu Jiangping^{**}, Huang Yuejin^{*}, Wang Xinlei^{***}, Zheng Aiwu^{****}, Shen Xi^{*}

(^{*} College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023)

(^{**} College of Education, Zhejiang University of Technology, Hangzhou 310023)

(^{***} Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, Urbana IL61801)

(^{****} JiaXiPera Compressor Co. Ltd, Jiaxing 314011)

Abstract

With the development of deep learning, more and more heat pump system fault diagnosis methods use deep learning technology and get good results. The fault diagnosis technology based on deep learning needs to rely on a large number of labeled fault data, but in reality, such data is very difficult to obtain, which limits the application of intelligent diagnosis technology. Aiming at this issue, the generative adversarial network (GAN) is proposed to learn the distribution of fault data and generate more labeled data to achieve the augmentation of the fault data set. For the complex operation data structure of the heat pump systems and the small difference of data value between different faults bring great difficulty to model learning, this paper proposes to use the heat pump system benchmark model to convert the operation data into residual data and use it as training data to reduce data complexity and increase the difference of data value. Using the MMD and 1-NN indicator to analyze the generated data, it is found that the distribution of the generated data is close to the real data, and the GAN model trained with residual data is of higher quality. Using the method of fault diagnosis to analyze the training results of models that use different amounts of generated data, it is found that the introduction of generated data can improve the accuracy of fault diagnosis under insufficient data conditions. The experimental results prove that the GAN-based data augmentation method can effectively reduce the dependence of intelligent diagnosis on labeled data, and has broad application prospects.

Key words: generative adversarial network (GAN), deep learning, data augmentation, heat pump system, fault diagnosis