

## 基于变分自编码器的时间序列生成及异常检测模型研究

郭昌林, 周在彦, 刘春秀, 李万彬, 刘奕敏, 李龙潭, 金桂玥, 薛志伟

(国网山东省电力公司德州供电公司, 山东 德州 253000)

**摘要:** 异常检测在工业设备故障监测等能源物联网场景中具有重要的应用价值, 能够帮助物联网系统实时识别时间序列中的异常模式, 从而提升系统的安全性、稳定性和运维效率。然而, 时序数据的稀缺性是制约模型性能的主要瓶颈之一, 这主要源于高质量标注的时序数据获取成本高昂, 以及工业生产等场景下的时序数据采集条件有限, 难以覆盖各种可能的情境。传统的数据增强方法难以有效捕获异常事件的复杂性和多样性, 进一步制约了检测模型的性能提升。为此, 提出了一种基于时间序列生成的数据增强方法, 以提高模型在数据稀缺条件下的异常检测能力。该方法利用变分自编码器生成模型, 在稀缺数据条件下合成具有真实性和多样性的时序数据, 从而有效缓解数据稀缺性对模型性能的限制, 显著提升异常检测的鲁棒性和准确性。实验结果表明, 该方法在智能制造等能源物联网场景中具有良好的适应性, 为构建高效、智能的异常检测系统提供了有效的技术支持。

**关键词:** 异常检测; 变分自编码器; 时间序列生成; 数据增强; 能源物联网

中图分类号: TN915.08

文献标志码: A

doi: 10.11959/j.issn.2096-3750.2026.00499

## Research on time series generation and anomaly detection model based on variational autoencoder

Guo Changlin, Zhou Zaiyan, Liu Chunxiu, Li Wanbin, Liu Yimin, Li Longtan, Jin Guiyue, Xue Zhiwei

State Grid Shandong Electric Power Company Dezhou Power Supply Company, Dezhou 253000, China

**Abstract:** Anomaly detection is proving to be a key capability in energy Internet of things(IoT) scenarios such as industrial equipment fault monitoring, enabling IoT systems to identify abnormal patterns in time-series data in real time, thereby improving system security, stability, and operational efficiency. However, the scarcity of time-series data is emerging as a major bottleneck limiting model performance, mainly due to the high cost of acquiring high-quality labeled data and the limited data collection conditions in industrial production, which make it difficult to cover all possible scenarios. Traditional data augmentation methods are increasingly being regarded as inadequate for capturing the complexity and diversity of abnormal events, further constraining the performance of detection models. To address these issues, a data augmentation method based on time-series generation was proposed to enhance the anomaly detection capability of models under data-scarce conditions. The method utilized a variational autoencoder generative model to synthesize realistic and diverse time-series data from limited samples, thereby mitigating the impact of data scarcity on model performance and significantly improving the robustness and accuracy of anomaly detection. Experimental results demonstrated that the proposed method exhibited good adaptability in energy IoT scenarios such as smart manufacturing, providing effective technical support for building efficient and intelligent anomaly detection systems.

**Key words:** anomaly detection, variational autoencoder, time series generation, data augmentation, energy Internet of things

收稿日期: 2025-06-10; 修回日期: 2026-02-05

通信作者: 周在彦, 601980730@qq.com

基金项目: 国网山东省电力公司科技项目 (No. 520608240006)

**Foundation Item:** State Grid Shandong Electric Power Company Science and Technology Project (No. 520608240006)

## 0 引言

随着能源物联网的发展,智能传感与数据分析技术在能源系统中的应用日益广泛。基于物联网技术,各类能源设备能够实时采集、传输并分析时序数据,从而实现智能监测与预测性维护。在此背景下,异常检测技术在工业生产、智能电网、配电系统等多个场景中扮演着关键角色。例如,在工业制造领域,部署于工厂的智能传感器可以实时监测关键设备(如涡轮机、泵、压缩机等)的振动、温度、电流等参数;异常检测技术能够及时识别设备运行异常并发出预警,从而有效预防故障、减少停机时间、提高设备维护效率。

然而,在能源物联网环境下,时序数据异常检测仍面临诸多挑战。一方面,高质量标注数据的获取成本高昂,尤其是在分布式能源系统中,数据采集受到物理环境和传感器分布的影响,数据稀缺性问题突出。另一方面,异常事件通常较为罕见,数据样本往往呈现不均衡性,导致传统异常检测模型难以学习有效的特征表示,进而影响检测的精度和召回率<sup>[1]</sup>。

针对上述挑战,本文提出一种基于生成式数据增强的异常检测方法,结合物联网环境中的多模态时序数据,以提升模型的检测性能。具体而言,本文采用变分自编码器技术,通过学习能源物联网传感器数据的模式,在数据稀缺条件下生成高质量的合成样本,进而增强异常检测模型的泛化能力。进一步,在数据生成过程中引入可控生成机制,使合成数据能够涵盖更多异常模式,为模型提供更具代表性和多样性的训练数据。实验结果表明,该方法在智能制造、工业设备预测性维护等应用场景中,能够有效提高异常检测的准确性和鲁棒性。

## 1 相关工作

### 1.1 时间序列生成技术

由于数据增强、异常检测、隐私保护和领域自适应等需求不断增长,时间序列生成(TSG, time series generation)作为一个新兴研究方向正受到日益广泛的关注<sup>[2-4]</sup>。TSG的目标是生成与原始时间序列相似的时间序列,在保留时间依赖性和维度相关性的同时,确保生成的时间序列对分类和预测等各种下游任务仍然有用。

近年来,时间序列生成方法主要围绕3种生成架构展开:生成对抗网络(GAN, generative adversarial network)、变分自动编码器(VAE, variational autoencoder)和扩散模型。其中,GAN在生成质量上表现出色,但其对抗性训练机制常伴随模型优化困难,训练过程容易不稳定甚至崩溃。相比之下,VAE则通过平衡生成数据的保真度与隐空间的统计一致性,进一步提升了模型的可解释性,使生成的时间序列数据更具现实意义。以分数扩散模型<sup>[5]</sup>和去噪扩散概率模型(DDPMs, denoising diffusion probabilistic models)<sup>[6]</sup>为代表的扩散模型,在生成质量上优于传统GAN,且避免了对抗性训练的优化难题,但目前相关研究较少。

典型的GAN模型包括一个生成器和一个鉴别器<sup>[7]</sup>。在训练过程中,生成器生成合成时间序列,鉴别器则试图将其与真实序列区分开来。这种对抗循环一直持续到鉴别器无法可靠地区分合成与真实数据。最终在生成阶段,只使用生成器来创建新的时间序列。

在时间序列生成任务中,许多模型基于生成对抗网络框架,并结合专门用于时序建模的神经网络结构,如循环神经网络(RNN, recurrent neural network)、长短期记忆网络(LSTM, long short-term memory)和Transformer,以捕捉时间序列数据中的时序和复杂依赖关系<sup>[8-10]</sup>。此外,部分研究通过设计新的度量准则或损失函数,使模型更好地适应特定时序特征<sup>[11-12]</sup>。也有一些方法通过增强传统GAN架构来改进效果,如添加额外的鉴别器、分类层、纠错机制或数据增强机制,从而生成具有特定时间属性的序列<sup>[13]</sup>。例如,TimeGAN融合自回归模型和对抗学习理念,共同学习时间序列的时间依赖性和潜在表示,以生成高质量的序列数据<sup>[14]</sup>。COSCI-GAN通过公共潜在空间生成各个时间序列,并利用中心判别器增强通道/特征间的动态相关性,以更好地保持数据内部结构<sup>[15]</sup>。AEC-GAN则引入误差校正模块,动态修正生成数据的偏差,并借助对抗样本增强训练,缓解分布偏移问题<sup>[16]</sup>。尽管基于GAN的模型在时间序列生成方面取得了较好效果,但其训练过程通常较为复杂,且需要大量计算资源与时间。尽管如此,GAN凭借其在生成质量上的表现,仍成为近年来生成合成数据的主要方法之一。

尽管GAN在生成质量方面具有一定优势，但其仍存在一定的局限性。一方面，GAN的训练过程往往不够稳定，且易出现模式崩溃现象，即生成器为欺骗鉴别器而反复输出一些相似且缺乏变化的序列。另一方面，GAN通常需要大量计算资源，属于资源密集型和任务密集型的任务。此外，GAN还可能过度拟合训练数据，从而限制生成时间序列的多样性。

变分自动编码器的基本模型包括编码器和解码器两部分<sup>[7]</sup>。编码器将输入时间序列转换为隐表示，通过捕获基本特征并生成均值、方差等统计参数，以建模时间序列固有的不确定性和变化性。解码器则负责从隐表示重构时间序列，从而保留模型再生时间模式的能力。训练阶段，解码器的目标是最大化重建损失，并减小隐表示分布与预设标准高斯分布之间的差异。生成阶段，则利用解码器从隐空间中采集并合成新的时间序列数据。

与GAN相比，尽管目前基于VAE的TSG方法研究较少，但它们有效地利用变分推理来捕获时间序列数据的复杂时序特征<sup>[17-19]</sup>。VAE的优势在于能够从输入数据中学习复杂的概率分布，这使它们在生成建模任务中具有良好效能。在许多应用场景下，VAE展现出比GAN更多的优势。通过将数据编码到一个较低维的隐空间中，VAE促进了有效的表示学习，这有助于理解和分离数据中变化的隐因素。此外，VAE可以使用基于梯度的优化技术进行训练，这不仅简化了参数迭代更新的过程，也使其能够在大规模数据集上实现高效扩展。

近年来，基于分数的扩散模型，特别是去噪扩散概率模型，由于在感知质量上优于GAN，同时避免了对抗性训练带来的优化挑战，已在图像、视频和文本生成等领域得到了广泛应用与研究<sup>[11,20-21]</sup>。扩散模型因其生成质量高、训练稳定、支持显式优化和细粒度控制，适合多种复杂任务，尤其在时间序列和特定条件生成方面展现出明显优势。扩散模型的实现过程主要包括前向扩散过程和反向去噪过程。在前向扩散过程中，扩散模型通过逐步向数据添加高斯噪声，将原始数据样本逐渐变得模糊，最终得到一个接近纯噪声的分布，该过程本质上是一个马尔可夫链。在生成阶段，扩散模型则利用训练好的神经网络逐步逆转前向扩散过程，从噪声数据中逐步恢复出接近原始数据的样本，这个过

程同样是一个马尔可夫链，用神经网络估计的去噪模型反向预测每一步的噪声样本。尽管扩散模型具备上述优点，但目前针对时间序列数据建模的研究仍然相对有限。

## 1.2 时序数据增强技术

自深度学习出现以来，科学界的大量研究和工都集中于解决和优化监督学习任务<sup>[1]</sup>。监督学习通常需要大规模、特征多样且标注完备的数据集，而随着当前模型规模的扩大，对训练数据量的需求也日益增长，获取符合要求的数据集变得越来越困难。通常，存在大量的公共存储库，从中可以获得适合大多数应用领域的训练数据集。尽管目前存在众多公共数据仓库，可为大多数应用领域提供训练数据集，但在时间序列领域，高质量数据集仍不易获得，普遍面临数据隐私困难、样本规模有限或类别不平衡等挑战。为此，数据增强（DA, data augmentation）技术成为应对数据稀缺问题的一种重要手段。

此外，时间序列数据具有特有的复杂性。不同时间序列的数据集可能具有不同的特征，因此，在增强过程中需要特别关注方法是否适配数据特性。为此，有必要分析和讨论各类时间序列数据增强技术，以更好地了解不同技术适用于何种数据类型。

通常，此类问题可以通过预处理数据集技术来解决，如子采样，或者在数据规模不足时，借助数据增强技术来提升数据的有效性和多样性<sup>[22-23]</sup>。

传统上，在图像识别等领域，常通过裁剪、缩放、镜像、颜色增强或翻译等变换来扩充数据。然而，由于时间序列数据分布的特殊性，这些方法不能直接应用于时间序列领域<sup>[24]</sup>。例如，如果要应用旋转来增强图像数据集，则可以旋转每个图像以生成新图像，但这不能直接在时间序列域中完成。如果一个时间序列样本被分成几个部分，并且这些部分使用它们之间的线性插值进行重组，则结果将无效，因为原始数据的趋势特征会因此被破坏。由于时间序列数据本身具有多样性，并非所有增强技术都适用于每个数据集。因此，一部分计算机视觉中的传统算法需经过调整才能适配时序数据，而在更多情况下，则必须设计新的特定算法来处理时间序列数据。

目前，时间序列数据增强的常用技术包括随机扰动、窗口切片等方法。在许多实际应用中，这些

技术通过扩大数据集、增加数据多样性,在提高模型的性能方面展现出良好的效果<sup>[3]</sup>。

然而,DA技术在时间序列数据上的应用仍面临一些挑战。首先,常见的DA方法(如简单扰动和数据切片)在扩充样本时,往往难以充分保持原始时间序列的时间依赖性和多维关联性;其次,生成数据的统计特性可能与原始数据存在明显偏差,导致生成样本的真实性不足,进而影响其在实际任务中的应用效果。此外,现有DA方法通常缺乏对特定情境或任务条件的适应能力,难以满足不同下游任务对数据特性的差异化需求。

为克服上述局限,研究者提出了TSG<sup>[2]</sup>。TSG的核心目标是生成与原始数据分布相似的时间序列数据,确保生成样本能够保留时间依赖性和多维关联性,以便在分类、异常检测等下游任务中具备实际应用价值。与单纯的DA不同,TSG不仅强调数据的数量扩展,还注重生成样本在时间模式、变化趋势、季节性等统计特性与真实数据的一致性,从而维持生成数据的真实性与实用性。基于此,TSG能够为不同下游任务提供更具针对性的合成数据,为解决数据稀缺问题提供了更为细致和有针对性的途径。

在许多时间序列应用场景中,TSG算法通过扩展数据集规模和提高样本多样性,有效提升了模型的泛化能力与稳健性。目前,这类算法已经在金融市场预测、医疗健康监测、工业设备故障检测等领域展现出显著成效并得到广泛应用<sup>[2-4]</sup>。此外,TSG也为应对数据隐私和安全性问题提供了新思路:其能够在保持数据原始特征的基础上生成合成数据,从而减少直接使用敏感数据的需求。因此,TSG作为一个快速发展的新兴研究方向,为时间序列数据的合成与增强提供了新的解决路径,并在异常检测、领域自适应、分类等下游任务中显现出明显的应用优势。

### 1.3 时间序列异常检测技术

时间序列是按时间顺序排列的一组数据点或观测值的集合,通常以等间隔方式采集。不同于普通数据,时间序列中各数据点之间并非相互独立,而是具有时序上的关联性,后续观测值往往受到先前观测的影响。

在时间序列中,存在的异常类型包括点异常、子序列异常以及时间序列分量异常<sup>[25]</sup>。点异常被定

义为时间序列中某个观测点的值不符合预期且出现较大偏差,其影响可能涉及单个或多个时间维度。子序列异常被定义为一段连续时间范围内的观测值整体行为存在异常,其影响范围可以是全局或局部的。时间序列分量异常被定义为时间序列数据中的趋势、季节性或其他结构性分量出现了不符合预期的变化。

根据其使用的技术不同,现有时间序列异常检测方法可划分为基于统计学的方法和基于机器学习的方法。

常见的基于统计学的异常检测方法包括Z分数、Grubbs假设检验、自回归整合移动平均(ARIMA, autoregressive integrated moving average)模型和基于经验累积分布函数(ECDF, empirical cumulative distribution function)的异常检测方法。

Z分数是一种常用的基于统计学方法,用于判断数据点与均值的偏离程度。当Z分数的绝对值超过设定的阈值,则认为该数据点为异常值。Z分数的优点是简单直观,适用于高斯分布的数据,但当数据具有强烈的趋势性或季节性时,可能不适用。

ARIMA模型是ARMA模型的扩展,专门用于处理非平稳时间序列。ARIMA模型通过加入差分成分来消除趋势性,使时间序列变得平稳。ARIMA模型的核心是自回归和移动平均成分,结合差分操作来建立时间序列模型。ARIMA模型适用于具有趋势性但没有季节性成分的时间序列。在异常检测中,ARIMA可用于对时间序列进行建模,并使用模型的预测值与实际观察值之间的残差来判断异常。当残差超出一定的范围时,该点被判定为异常。

经验累积分布函数是一种统计工具,用于估计样本数据的累积分布情况。通过计算数据点的累积概率,可以判断其是否显著偏离数据整体分布,从而识别异常点。对于一个给定的样本数据集 $X = (x_1, \dots, x_n)$ ,经验累积分布函数 $F_n(x)$ 可表示为:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n 1(x_i \leq x) \quad (1)$$

其中, $n$ 表示样本数量; $1(x_i \leq x)$ 表示指示变量,当 $x_i \leq x$ 时,取值为1,其余为0; $F_n(x)$ 表示数据点 $x$ 在样本分布中的累积概率。如果 $F_n(x)$ 取值很小,表明该点在样本分布中显著偏离,可能是异常点。利用这些概率值可计算出每个数据点的异常分

数，从而实现对数据点的排序和筛选。

传统基于机器学习的方法通过学习时间序列数据的特征来构建模型，进而识别异常值。这类方法广泛应用于时间序列异常检测领域，其常见类型包括基于距离、基于聚类、基于分类等方法<sup>[26]</sup>。

基于距离的异常检测方法通过度量数据点之间的距离来识别异常点。这类方法假设正常的点通常与其他点的距离较小，而异常点往往距离其他点较远。 $k$ 近邻算法 ( $k$ NN,  $k$ -nearest neighbors algorithm) 通过计算每个数据点与其最邻近点的距离，将那些与其他点距离较远的点标记为异常值。 $k$ NN方法直观且易于实现，但在处理高维数据时，计算复杂度较高。

基于聚类的异常检测方法将数据划分为多个簇，并根据簇的大小和形状来识别异常点。聚类方法假设正常点倾向于形成密集的簇，而异常点则往往孤立或远离主要簇。 $k$ 均值聚类通过最小化簇内平方误差将数据点分组，位于簇外的数据点被认定为异常。然而， $k$ 均值对初始中心的选择敏感，并且可能会忽略具有不同密度的异常点。

基于分类的异常检测方法将正常数据与异常数据区分开来，常见的分类方法包括支持向量机和孤立森林 (i Forest, isolation forest)。支持向量机通过寻找一个最优的超平面来区分正常数据和异常数据，具备处理高维数据的能力，并可通过核函数提高模型的非线性表达能力。iForest认为在数据集中的某个点如果更容易与其他点分开 (即“孤立”)，那么这个点更有可能是异常点。异常点通常在特征空间中处于稀疏位置，而正常点倾向于聚集。iForest通过随机选择特征和特征的分裂值构建二叉树，分裂过程相当于将数据不断切割。由于异常点距离其他点较远或分布稀疏，通常需要更少的分裂操作就能将其孤立。尽管iForest在思路上具有分类的特点，但由于其无监督的特性，与传统监督分类方法 (如支持向量机、决策树等) 有所不同。它特别适合高维数据和复杂的时间序列异常检测任务，具有高效和鲁棒的特点。

## 2 模型

### 2.1 问题定义

时间序列生成定义：给定一个时间序列数据集  $T$ ， $T$  由  $N$  个长度为  $L$  的时间序列构成，即  $T =$

$(s_1, \dots, s_N)^T$ ，其中每个时间序列  $s_i$  可以视为一个维度为  $L$  的向量，即  $s_i = (x_{i,1}, \dots, x_{i,L})$ ， $x_{i,j}$  是时间序列  $s_i$  对应于时间点  $t_j$  时的数据。不妨设  $p(s_1, \dots, s_N)$  是给定时间序列  $T$  的真实分布。可控时间序列生成的目标是创建一个合成时间序列  $T^{\text{gen}} = (s_1^{\text{gen}}, \dots, s_N^{\text{gen}})^T$ ，使得该序列的分布  $q(s_1^{\text{gen}}, \dots, s_N^{\text{gen}})$  与真实分布  $p(s_1, \dots, s_N)$  接近，从而  $T^{\text{gen}}$  与  $T$  能表现出一致的统计特性。

本文中时序数据生成的目标是根据已有时间序列数据，生成多样化的时序数据，帮助模型识别和检测各种类型的异常，进而提高模型在异常数据稀缺或不均衡的情况下检测异常的能力。

### 2.2 模型框架

模型框架由四部分组成，包括预处理、模型训练、数据增强和异常检测。首先，将时序数据集进行预处理，包括归一化、填充缺失值、对齐时序分辨率等；然后，将处理好的统一化时序数据送入生成模型中，进行训练生成；最后，将生成数据与原始数据通过数据增强送入异常检测模型，达到提高异常检测模型性能的效果。时间序列生成异常检测框架如图1所示。

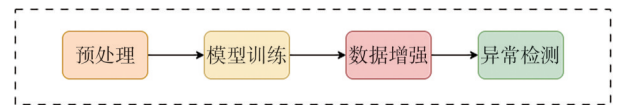


图1 时间序列生成异常检测框架

#### 2.2.1 预处理

在实际应用中，时序数据往往来自多个传感器或数据源，不同数据通常具有不同的模态 (如电流、电压、温度等)。为确保模型能够有效处理这些多模态数据，需要对其进行统一预处理，具体包括以下几个方面。

(1) 归一化：不同模态的数据往往具有不同的量纲与数值范围，采用归一化方法 (如最小最大归一化) 处理可消除不同特征之间的尺度差异。

(2) 缺失值填充：多模态数据在采集过程中可能出现缺失值。常见的填充方法有均值填充、插值法。本文采用均值填充进行缺失值处理。

(3) 时序分辨率对齐：不同来源的数据往往具有不同的时间分辨率 (即时间戳之间的间隔可能不同)，为保证模型能够有效学习时序依赖关系，需要对不同时间分辨率的数据进行对齐。根据实际任务需要，可将数据划分为等长或不等长的时间窗

口，以便模型更好地捕捉时序局部模式与长期依赖。

### 2.2.2 TimeVAE 训练

TimeVAE (time-series variational autoencoder) 是一种融合变分自编码器与时间序列数据特性的生成模型<sup>[7]</sup>，适用于生成具有时序依赖性的合成数据。其编码器部分用于从输入的时间序列数据中提取隐变量。TimeVAE 的编码器通常由卷积神经网络 (CNN, convolution neural network) 和全连接层构成。编码器的输出为隐变量的均值和方差，这些变量共同构成了数据的低维隐空间表示。TimeVAE 结构如图 2 所示。

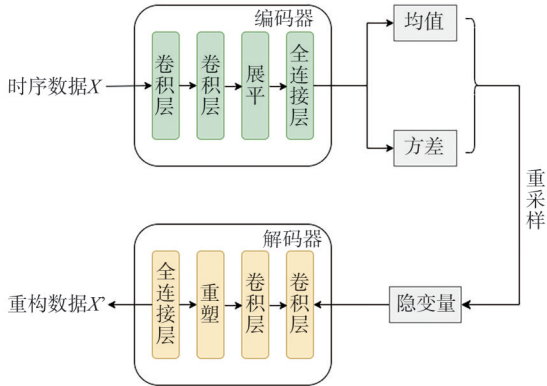


图 2 TimeVAE 结构

VAE 模型通过重参数化技巧，允许从隐空间采样，进而生成新的时间序列数据。隐空间的设计通常考虑了时间序列数据的时序特征，以保持采样过程中的时序一致性。

生成隐空间变量后，解码器则负责将隐变量转换回原始时间序列的重建值，以有效捕捉时间序列中的长期依赖关系，解码器同样由 CNN 构成。在多层卷积操作后，特征被送入全连接层进行整合。最终，解码器输出一个与输入序列相同维度的重建序列，用来与真实数据进行对比，计算重建误差。

$$L_{\theta, \phi} = -E_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)] + D_{KL}(q_{\phi}(z|x) \parallel p_{\theta}(z)) \quad (2)$$

TimeVAE 的损失函数由重建误差和 KL 散度两部分组成。重建误差用于度量生成数据与真实数据之间的差异，其优化目标是最小化该误差，以提升数据的重建质量。KL 散度则用于约束隐变量的分布，使其接近于先验分布，从而在维持数据表征能力的同时，增强隐空间的结构化程度，防止模型过拟合。TimeVAE 通过最大化变分下界 (ELBO, evi-

dence lower bound) 进行训练，使得隐空间能够有效捕捉数据中的时序模式，并优化生成能力。模型一旦训练完成，便可从隐空间中采样生成新的时间序列数据，用于数据增强或模拟潜在模式。

### 2.2.3 数据增强

数据增强是通过生成新的数据样本来扩展现有数据集的方法，旨在提升模型的鲁棒性和泛化能力。在时间序列异常检测任务中，数据增强的目标是通过合成异常数据并调整其与正常数据的比例，缓解数据类别不平衡问题，从而增强模型对不同异常模式的识别能力。

常见的时间序列数据增强策略包括随机扰动、窗口切片等。这些算法通过扩大数据集来提高模型的性能，且在许多时间序列应用中显示出良好的效果。在本研究中，提出使用基于 TimeVAE 的时间序列生成模型进行数据增强，以进一步提升异常检测的效果。

在异常检测任务中，异常数据通常稀缺。因此，首先利用生成模型生成具有多样化和代表性的数据，以扩充数据。生成的数据应具备与正常数据不同的特征，以增加数据多样性，从而使检测模型能够识别潜在的异常模式。

将生成的异常数据与原始正常数据混合是数据增强过程的关键步骤。本文采用以下两种数据混合策略。

(1) 随机混合：从原始正常数据中随机选择一部分样本，与生成的异常数据按一定比例混合。该方法有助于平衡数据集中正常与异常数据的比例，避免模型对正常数据的过拟合。

(2) 加权混合：根据一定的加权策略，将正常数据与异常数据按预设的比例进行融合。该策略允许根据任务需求灵活调整生成数据和正常数据的比例，从而增强模型对不同异常情况的适应能力。

通过将正常数据和生成的异常数据结合，构建出一个增强后的训练集。该增强数据集被用于训练异常检测模型，以帮助模型更准确地学习到正常与异常模式的区别。

### 2.2.4 异常检测

在本文中，选用 ECOD (empirical cumulative distribution-based outlier detection) 和 iForest 作为异常检测模型。这两种方法均为无监督异常检测的经典算法，具有高效性和良好的泛化能力，可从不同

角度识别时间序列数据中的异常模式：ECOD基于统计特征与数据分布特性进行分析，而iForest则依托随机化的树结构捕捉异常。

ECOD是一种基于经验累积分布函数的无监督异常检测方法<sup>[26]</sup>。其核心思想是通过每个特征的累积分布进行建模，从而量化样本的异常程度。ECOD逐特征计算累积分布函数，避免了复杂的参数调整，运行效率高，且不依赖任何特定的分布假设，适用于任意数据集。

iForest是一种基于决策树的无监督异常检测方法<sup>[27]</sup>，通过构建多棵随机化的二叉树（隔离树）来对样本的异常性进行建模。iForest通过随机选择特征和分割点，构建隔离树，递归划分数据样本。异常点由于特征值较为稀疏，通常能在较少划分步骤中被“孤立”，因此其路径长度较短。iForest根据样本所在树种的平均路径长度计算异常得分，得分越高表明样本越可能是异常点。iForest具有近似线性的时间复杂度，适用于大规模高维数据，且不依赖特定分布，适应性强，能够应对不同数据集的异常检测任务。

### 3 实验

#### 3.1 数据集

为评估异常检测方法的性能，本文采用UCR时间序列数据集进行实验。UCR数据集是一个广泛使用的标准化时间序列数据集集合，包含多个领域的多样化时间序列数据。该集合中的每个数据集提供独立的训练集和测试集，其数据均为单变量时间序列，各数据集的序列长度和类别数量不尽相同，故而对时序生成任务及检测任务具有显著的代表性和挑战性。在能源物联网环境中，智能传感器能够实时采集各类设备的运行状态数据，并通过物联网通信协议将其传输至边缘或云端进行分析。异常检测在能源物联网场景中至关重要，可用于预测设备故障、优化维护策略、提高系统安全性和运营效率。为确保实验数据与能源物联网应用场景的高度匹配，在从UCR数据集中挑选具有代表性的时间序列数据集时，需重点关注以下方面。

(1) 所选数据集应涵盖能源物联网中的典型时间序列模式，如工业设备状态监测、飞行器轨迹跟踪、制造过程质量控制等。

(2) 数据中需包含可辨识的异常模式，如设备

故障、运行偏差、能耗异常或环境干扰等，以评估方法在实际场景中的有效性。

(3) 数据集应具备足够的复杂度，能够反映能源物联网场景中常见的高维度、非平稳或噪声干扰特性，以验证检测方法的鲁棒性。

针对异常检测任务，本研究从UCR数据集中挑选了多个满足标准的数据集，这些数据集涵盖了物联网场景中的时间序列模式和异常类型。

(1) 工业传感器数据（Car和FreezerRegular-Train）：该数据集包含工业设备（如风机、电机、压缩机等）的传感器时间序列数据，包括温度、振动、能耗、电流等运行参数。异常可能表现为运行状态不稳定、传感器读数异常等。在能源物联网环境中，智能传感器可实时采集设备状态，并上传至物联网平台。基于此类数据集的异常检测可用于预测设备故障、优化维护策略、减少非计划停机时间，从而提高设备能效和系统稳定性。

(2) 运动轨迹数据（如Plane）：该数据集存储了飞行器的航迹时间序列数据，包括位置、速度、加速度等运动参数。异常可能表现为飞行姿态异常、失速等。在新能源基础设施巡检（如风电场、输电线路等）中，无人机常被用于自动巡检任务。异常检测有助于及时发现无人机运行异常，保障巡检数据的可靠性，进而提高新能源设施的安全性和巡检效率。

(3) 生产质量监控数据（如Wafer）：该数据集来源于半导体制造过程，记录了晶圆的生产质量时间序列数据，包括传感器测量的电流、电压、温度等工艺参数。异常可能表现为某些关键工艺参数偏离正常范围，导致产品质量缺陷。在智能制造及半导体生产中，能源物联网系统可实时监测生产设备的运行状态和能耗情况。基于该数据集的异常检测有助于识别生产异常、优化工艺参数、降低能耗浪费，提高制造质量和生产效率。

数据集概况见表1。表1展示了本研究选取的6个数据集中训练集和测试集的样本数量，以及每个数据集的时间序列长度。为方便表示，在后续的内容中，将采用缩写表示，使用FRT表示Freezer Regular Train数据集，使用IWbS表示Insect Wingbeat Sound数据集，使用IPD表示Italy Power Demand数据集。

为适应异常检测任务，对数据进行了预处理，

表1 数据集概况

数据集	训练集样本数量	测试集样本数量	时间序列长度
Car	49	60	577
Freezer Regular Train	75	2 850	301
Insect Wingbeat Sound	200	1 980	601
Italy Power Demand	34	1 029	24
Plane	96	105	144
Wafer	903	6 164	152

并根据任务需求,重新定义异常标签。首先,统计数据集中各类别样本的数量;然后,将样本量最少的类别定义为异常类别,并将这些样本标注为异常标签。这一定义方法基于异常检测中的常见假设,即异常数据在分布上通常是稀有的或不平衡的。其余样本类别被统一定义为正常类别,并标注为正常标签。这样做不仅符合异常检测任务的实际需求,而且确保了样本标注的合理性,使模型能够专注于捕捉小类别样本的独特模式。

### 3.2 基线模型

在评估数据增强方法对异常检测性能的影响时,基线模型的选取具有重要意义。例如,在本研究中,选取基于不同数据增强技术的异常检测模型作为基线,具体包括:基于采样的窗口滑动(Window Slicing)、基于变形的窗口扭曲(Window Warping),以及基于GAN的TimeGAN<sup>[28]</sup>。

窗口滑动技术指在时间序列数据上设置一个固定大小的滑动窗口,该窗口沿时间轴逐步移动,从而生成一系列具有固定长度的子序列。该方法广泛应用于特征提取、信号处理及时间序列预测等任务,能够有效捕获序列中的局部时间依赖和动态变化特征。

设时间序列数据为 $X = (x_1, \dots, x_n)$ ,其中, $n$ 为时间序列的总长度。滑动窗口技术通过一个固定长度为 $w$ 的窗口在时间轴上移动,生成一系列子序列 $S = \{S_1, \dots, S_m\}$ ,其中 $m = n - w + 1$ 是滑动窗口的生成数量。 $S_i$ 表示如下:

$$S_i = (x_i, x_{i+1}, \dots, x_{i+w-1}), i \in \{1, 2, \dots, m\} \quad (3)$$

其中,窗口起始位置从序列的第一个点 $x_1$ 开始,按照步长 $s$ (通常 $s = 1$ )逐步向后滑动。

窗口扭曲技术则是在时间序列数据处理中对窗口进行灵活调整,包括形变、扩展或缩放,以适应时间序列的多样性特征或捕获关键动态模式。通过对窗口位置或大小进行规则化调整,该方法能够更

加充分地表征数据的时变特性和局部动态行为,从而为后续异常检测任务提供更为丰富的特征信息。

设时间序列数据为 $X = (x_1, \dots, x_n)$ 其中, $n$ 为时间序列的总长度。窗口的长度为动态变量 $w_i$ ,窗口扭曲生成的子序列 $S = \{S_1, \dots, S_k\}$ 可表示为:

$$S_i = (x_{p_i}, x_{p_i+1}, \dots, x_{p_i+w_i-1}), i \in \{1, 2, \dots, k\} \quad (4)$$

其中, $p_i$ 表示窗口的起始位置,第 $i$ 个窗口可以根据规则(如特征突变点或固定步长)动态选取起始点; $w_i$ 表示第 $i$ 个窗口的长度,可以根据序列变化特性(如能量或振幅)进行动态调整。

TimeGAN是一种基于GAN的针对时间序列数据设计的生成模型,能够同时学习时间序列的时间依赖关系和潜在数据分布。TimeGAN由自编码器、生成器和判别器构成,结合对抗学习和自监督学习来生成高质量的合成时间序列。其核心优势在于能够保持时间序列的动态特征,适用于不同领域的数据增强和模拟仿真任务。

在实验设置方面,针对TimeVAE模型设置编码器与解码器的CNN结构为2层CNN,通道数分别为32与64,隐变量维度为16,训练轮数为100轮,优化器选用Adam。对于TimeGAN模型,其自编码器部分采用RNN结构,隐变量维度为12,训练轮数为500轮,同样选用Adam优化器进行训练。

### 3.3 评估标准

在异常检测任务中,模型性能的评估主要围绕增强数据集上的检测效果展开。具体采用以下指标。

(1) 准确率 (Accuracy): 用于衡量模型整体预测性能,表示模型正确分类的样本占总样本的比例。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (5)$$

其中,TP (true positive) 为模型正确检测出的异常样本数, TN (true negative) 为未被模型检测出的正常样本数, FP (false positive) 为模型错误检测为异常的正常样本数, FN (false negative) 为实际异常但未被模型检测出的样本数。

(2) 精确率 (Precision): 反映模型检测出的异常中实际为异常的比例。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

(3) 召回率 (Recall): 反映模型从实际异常样本中成功检测出的比例。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

(4) F1 分数 (F1-score)：综合精确度与召回率的加权平均值，是全面评估模型性能指标。

$$\text{F1 - score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

在异常检测任务中，数据通常存在类别不平衡，单独使用准确度可能会掩盖模型在异常检测上的不足。因此，需要结合精确率、召回率、F1 分数等多维度指标，从不同角度全面评估模型性能。通过这些指标的综合分析，能够更准确、全面地量化数据增强技术在异常检测中的有效性。

### 3.4 实验结果

在UCR中的6个数据集上，对不同数据增强方法在ECOD和iForest两个异常检测模型中的效果进行验证，不同指标的对比结果分别如图3~图10所示。

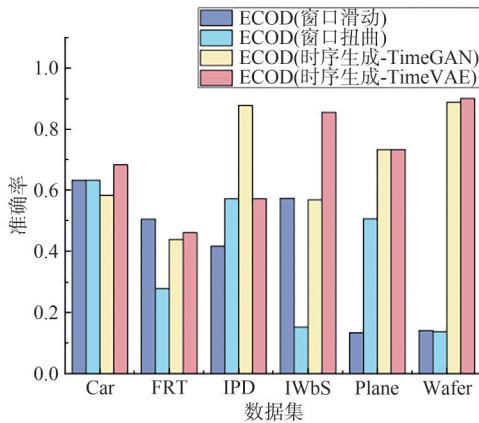


图3 不同数据增强方法经ECOD异常检测后的准确率对比

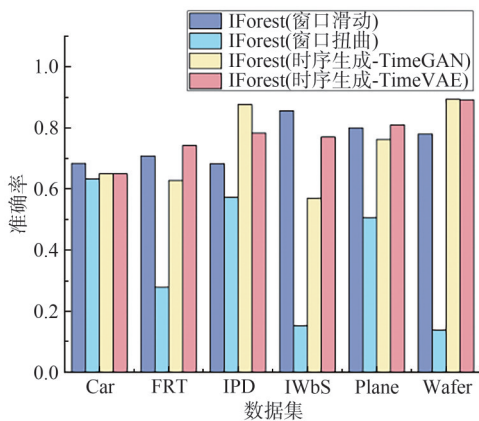


图4 不同数据增强方法经iForest异常检测后的准确率对比

从图3和图4来看，基于时序生成的数据增强方法在大多数数据集上的准确率均优于窗口滑动和

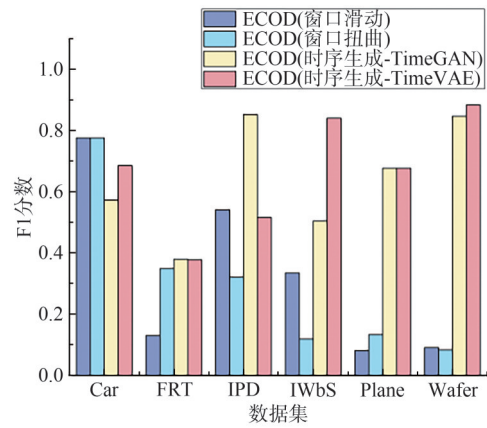


图5 不同数据增强方法经ECOD异常检测后F1分数对比

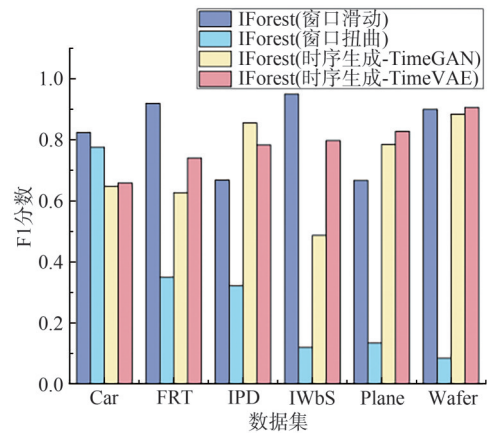


图6 不同数据增强方法经iForest异常检测后F1分数对比

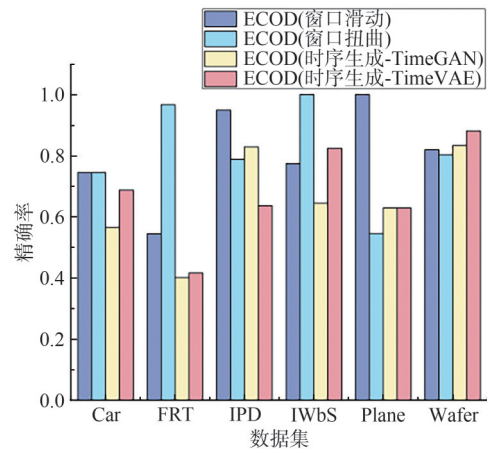


图7 不同数据增强方法经ECOD异常检测后精确率对比

窗口扭曲的数据增强方法。具体而言，在ECOD模型中，使用TimeVAE进行数据增强后，模型的准确率比窗口滑动方法提升约178%，比窗口扭曲方法提升189%；在iForest模型上，使用TimeVAE进行数据增强后，其准确率比窗口扭曲方法提升203%。对比使用两种时序生成增强方法的模型准

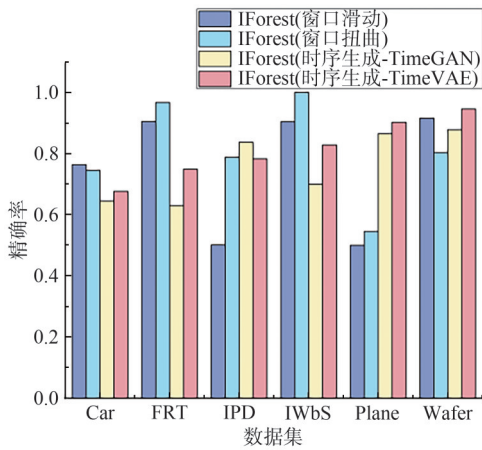


图8 不同数据增强方法经iForest异常检测后精确率对比

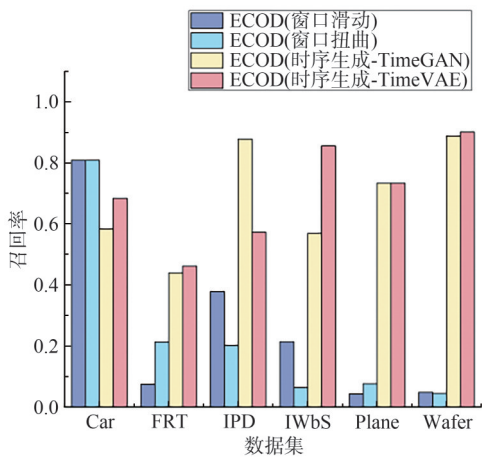


图9 不同数据增强方法经ECOD异常检测后召回率对比

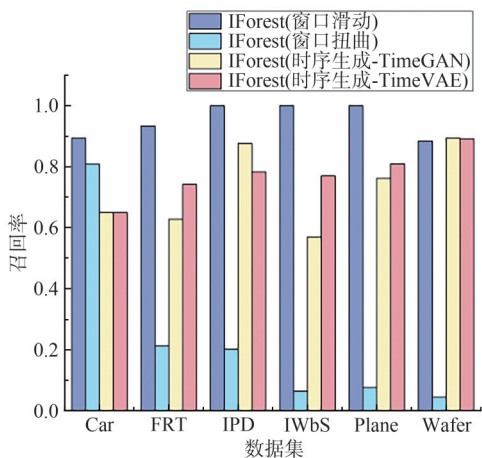


图10 不同数据增强方法经iForest异常检测后召回率对比

精确率可知，在 ECOD 模型中，除 IPD 数据集，TimeVAE 在其余数据集上的表现均优于 TimeGAN；在 iForest 模型上，TimeVAE 在 Car、FRT、IWbS 和 Plane 数据集上的表现均优于 TimeGAN。

从图 5 和图 6 来看，基于时序生成的数据增强

方法显著提升了模型的 *F1* 分数。在 ECOD 模型上，使用 TimeVAE 进行数据增强后，模型的 *F1* 分数比使用窗口滑动的数据增强方法提升了 322%，比使用窗口扭曲的数据增强方法提升了 338%，表明基于时序生成的数据增强方法更适合处理复杂的时间依赖关系。特别是在复杂数据集中，基于时序生成的数据增强方法优势尤为突出。例如，在 Wafer 数据集上，ECOD 模型的 *F1* 分数从 0.09 提升到 0.88，iForest 模型的 *F1* 分数从 0.08 提升到 0.90。对比使用两种不同时序生成增强方法的模型 *F1* 分数可以看到，除在 IPD 数据集上 TimeGAN 的表现优异外，在其余数据集上 TimeVAE 的 *F1* 分数更高，这说明 TimeVAE 在不同数据集上适用性更高、鲁棒性更强、学习效果更好。

从图 7 和图 8 来看，在大多数数据集（如 FRT、IPD、IWbS、Plane、Wafer）上，基于时序生成的数据增强方法相比窗口滑动方法显著提升了检测性能，特别是在样本量大、时间序列长度长的复杂数据集中，基于时序生成的数据增强方法表现更为优异。例如，在 Wafer 数据集中，ECOD 使用 TimeVAE 时序增强后精确率从 0.82 提升到 0.88，iForest 的精确率从 0.92 提升到 0.95。同时，与窗口滑动和窗口扭曲方法相比，基于时序生成的数据增强方法更具稳定性。对比使用两种不同时序生成方法的模型精确率可以看到，除了在 IPD 数据集上 TimeGAN 的精确率高于 TimeVAE 外，在其余数据集上 TimeVAE 的精确率均高于 TimeGAN。

从图 9 和图 10 来看，基于时序生成的数据增强方法在召回率上表现稳定，在 ECOD 和 iForest 模型上的平均召回率分别达到 0.70 和 0.77，均表现出良好的检测效果。窗口扭曲方法只在 Car 数据集上有一定效果，在其余数据集上表现较差；窗口滑动方法在 iForest 上呈现出优异的效果，但 ECOD 上表现不佳，平均召回率只有 0.26。对比两种基于时序生成的数据增强方法可以发现，除了 IPD 数据集外，TimeVAE 在其余所有数据集上的召回率均优于 TimeGAN，进一步验证了 TimeVAE 在异常检测任务中的稳定性和综合优势。

总的来说，在同一异常检测模型中，TimeVAE 在异常检测任务中取得了更好的检测效果和更高的准确率。TimeVAE 在多个数据集上均取得了显著的性能提升，表明该方法不仅适用于特定类型的时间

序列，还具备较强的泛化能力。相比窗口滑动和窗口扭曲方法，TimeVAE能更好地适配不同数据集的复杂时序特性，带来更为稳定的性能表现。与TimeGAN相比，TimeVAE在不同数据集上的检测效果更加优异，展现出更强的鲁棒性和泛化能力。时序生成方法通过增强数据的时间依赖关系，帮助模型更精准地识别异常行为，在复杂数据集上的表现尤为突出，进一步验证了该方法在复杂时序任务中的优越性。

#### 4 结束语

本文提出了一种基于生成技术的时序数据增强方法，在能源物联网场景下，尤其在数据稀缺的异常检测任务中，展现了显著的优势。该方法利用变分自编码器等生成模型，在稀缺数据条件下合成兼具真实性和多样性的时序数据，从而有效弥补真实数据的不足。这些生成的时序数据不仅丰富了物联网系统的训练数据集，还提升了模型在异常检测任务中的鲁棒性和准确性。

在实际工业物联网应用中，本文方法可广泛适用于电力设备监测、智能制造、智慧城市等场景。例如，在电力设备故障检测中，传感器采集的运行数据往往包含大量正常工况，而异常样本稀缺。通过本方法，可以在已有数据基础上生成多样化的故障模式，从而提高故障诊断系统对新型异常的检测能力。在智能制造中，不同生产批次和工艺流程可能导致数据分布的变化，本文方法能够增强模型对环境变化的适应性，减少对昂贵人工标注的依赖。

此外，为了适配不同工业场景，本文方法可结合特定场景的数据分布特征，调整先验分布，以确保生成数据符合行业规范。例如，在电力系统中可结合电力负荷曲线的统计特性，在智能制造中可引入工艺参数的约束条件，从而提升生成数据的实用性和可信度。

总之，该方法不仅能弥补数据稀缺问题，还能通过生成多样化的时序数据，使模型学习到更广泛的运行模式，即使在面对未知异常时仍能保持较高的检测性能，从而为工业物联网的智能化运维提供强有力的技术支撑。尽管本文所提方法在多个方面取得了显著成果，但仍存在一些局限性，未来的工作可以从以下几个方面进行改进<sup>[29-30]</sup>。

(1) 对特定条件的敏感性：生成模型的性能很大程度上依赖于输入数据的质量。如果原始数据存在大量的噪声或异常，生成的数据也可能受到负面影响。不同生成模型在不同任务中的表现可能有所差异，选择合适的生成模型对最终效果至关重要。

(2) 计算资源需求：生成模型的训练过程通常需要较大计算开销，特别是在大规模数据集上，需要充足的计算资源支持。在实际应用中，生成模型的推理速度也是一个重要的考量因素，尤其是在需要实时处理的工业场景中。

#### 参考文献：

- [1] Wang Q Y, Farahat A, Gupta C, et al. Deep time series models for scarce data[J]. *Neurocomputing*, 2021, 456:504-518.
- [2] Iglesias G, Talavera E, González-Prieto Á, et al. Data augmentation techniques in time series domain: a survey and taxonomy[J]. *Neural Computing and Applications*, 2023, 35(14): 10123-10145.
- [3] Navidan H, Moshiri P F, Nabati M, et al. Generative adversarial networks (gans) in networking: a comprehensive survey & evaluation[J]. *Computer Networks*, 2021, 194:108149.
- [4] Rigaki M, Garcia S. Bringing a gan to a knife-fight: adapting malware communication to avoid detection[C]//*Proceedings of the 2018 IEEE Security and Privacy Workshops (SPW)*. Piscataway: IEEE Press, 2018:70-75.
- [5] Song Y, Sohl-Dickstein J, Kingma D P, et al. Score-based generative modeling through stochastic differential equations[PP]. V2. arXiv(2021-02-10)[2025-06-10]. arXiv: 2011.13456.
- [6] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. *Advances in neural information processing systems*, 2020, 33: 6840-6851.
- [7] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. *Advances in neural information processing systems*, 2014, 27.
- [8] Ding D Z, Zhang M, Huang Y M, et al. Towards backdoor attack on deep learning based time series classification[C]//*Proceedings of the 2022 IEEE 38th International Conference on Data Engineering (ICDE)*. Piscataway: IEEE Press, 2022:1274-1287.
- [9] Li Z J, Tang Y C, Li W, et al. Learning disentangled representation with pairwise independence[C]//*Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence (AAAI'19/IAAI'19/EAAI'19)*. Palo Alto, CA: AAAI Press, 2019: 4245-4252.
- [10] Locatello F, Bauer S, Lucic M, et al. Challenging common as-

- sumptions in the unsupervised learning of disentangled representations[C]//Proceedings of the 36th International Conference on Machine Learning. Brookline, MA: PMLR, 2019: 4114-4124.
- [11] Locatello F, Poole B, Rätsch G, et al. Weakly-supervised disentanglement without compromises[C]//Proceedings of the 37th International Conference on Machine Learning. Brookline, MA: JMLR, 2020:6348-6359.
- [12] Locatello F, Tschannen M, Bauer S, et al. Disentangling factors of variations using few labels[PP]. V2. arXiv (2020-02-14)[2025-06-10]. arXiv: 1905.01258.
- [13] Alcaraz J M L, Strodthoff N. Diffusion-based time series imputation and forecasting with structured state space models[PP]. V1. arXiv(2022-08-19)[2025-06-10]. arXiv: 2208.09399.
- [14] Yoon J, Jarrett D, Van der Schaar M. Time-series generative adversarial networks[C]// Proceedings of the 33rd International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2019:5508-5518.
- [15] Seyfi A, Rajotte J, Ng R T. Generating multivariate time series with common source coordinated GAN (COSCI-GAN) [C]//Proceedings of the 36th International Conference on Neural Information Processing System. New York: Curran Associates Inc., 2022: 32777-32788.
- [16] Wang L, Zeng L, Li J. AEC-GAN: adversarial error correction GANs for auto-regressive long time-series generation[J].Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(8): 10140-10148.
- [17] Desai A, Freeman C, Wang Z H. et al. TimeVAE: a variational auto-encoder for multivariate time series generation[PP]. V1. arXiv(2021-11-15)[2025-06-10]. arXiv: 2111.08095.
- [18] Lee D, Malacarne S, Aune E. Vector quantized time series generation with a bidirectional prior model[C]//Proceedings of The 26th International Conference on Artificial Intelligence and Statistics. Brookline, MA, USA: PMLR, 2023:7665-7693.
- [19] Li H M, Yu S J, Principe J, et al. Causal recurrent variational auto-encoder for medical time series generation[C]//Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence. Palo Alto, CA: AAAI Press, 2023: 8562-8570.
- [20] Saharia C, Chan W, Chang H W, et al. Palette: image-to-image diffusion models[C]//Proceedings of the ACM SIGGRAPH 2022 Conference Proceedings (SIGGRAPH '22). New York, ACM Press: 2022:1-10.
- [21] Dhariwal P, Nichol A Q. Diffusion models beat gans on image synthesis[C]//Proceedings of the 35th International Conference on Neural Information Processing System. New York: Curran Associates Inc.,2021:8780-879.
- [22] Duong H-T, Nguyen-Thi T-A. A review: preprocessing techniques and data augmentation for sentiment analysis[J]. Computational Social Networks, 2021,8:1.
- [23] Felix E A, Lee S P. Systematic literature review of preprocessing techniques for imbalanced data[J]. IET Software, 2019, 13(6): 479-496.
- [24] Mikołajczyk A, Grochowski M. Data augmentation for improving deep learning in image classification problem[C]//Proceedings of the 2018 International Interdisciplinary PhD Workshop (IIPhDW). Piscataway: IEEE Press, 2018: 117-122.
- [25] Blázquez-García A, Conde A, Mori U, et al. A review on outlier/anomaly detection in time series data[J]. ACM Computing Surveys (CSUR), 2021, 54(3): 1-33.
- [26] Nassif A B, Talib M A, Nasir Q, et al. Machine learning for anomaly detection: a systematic review[J]. IEEE Access, 2021, 9: 78658-78700.
- [27] Li Z, Zhao Y, Hu X Y, et al. ECOD: unsupervised outlier detection using empirical cumulative distribution functions[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(12): 12181-12193.
- [28] Xu D, Wang Y J, Meng Y L, et al. An improved data anomaly detection method based on isolation forest[C]//Proceedings of the 2017 10th international symposium on computational intelligence and design (ISCID). Piscataway: IEEE Press, 2017: 287-291.
- [29] Rashid K M, Louis J. Times-series data augmentation and deep learning for construction equipment activity recognition[J]. Advanced Engineering Informatics, 2019, 42: 100944.
- [30] Wen Q S, Sun L, Yang F, et al. Time series data augmentation for deep learning: a survey[PP]. V4. arXiv (2022-03-31)[2025-06-10]. arXiv: 2022.12478.

## [作者简介]



郭昌林(1982-), 男, 国网山东省电力公司德州供电公司高级工程师, 主要研究方向为电力系统及其自动化。



周在彦(1992-), 男, 国网山东省电力公司德州供电公司工程师, 主要研究方向为配网协同调度。



刘春秀(1975-), 女, 国网山东省电力公司德州供电公司正高级工程师, 主要研究方向为电力系统及其自动化。



李龙潭(1987-), 男, 国网山东省电力公司德州供电公司工程师, 主要研究方向为电网自动化。



李万彬(1974-), 男, 国网山东省电力公司德州供电公司高级工程师, 主要研究方向为电力系统及其自动化。



金桂玥(1990-), 女, 国网山东省电力公司德州供电公司工程师, 主要研究方向为带电作业。



刘奕敏(1998-), 女, 国网山东省电力公司德州供电公司助理工程师, 主要研究方向为配电网自动化。



薛志伟(1986-), 男, 国网山东省电力公司德州供电公司高级工程师, 主要研究方向为配网不停电作业。