Computer Science and Application 计算机科学与应用, 2026, 16(11), 117-129
Published Online November 2026 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2026.1611330

Hans汉斯

基于条件扩散噪声预测的锂离子电池车辆级故障检测

刘　红[1,2]

[1]安徽理工大学计算机科学与工程学院，安徽　淮南
[2]合肥综合性国家科学中心能源研究院(安徽省能源实验室)，安徽　合肥

收稿日期：2026年10月5日；录用日期：2026年11月4日；发布日期：2026年11月12日

摘　要

针对真实车辆充电数据中故障样本稀缺，以及电压和温度响应随电流和荷电状态等运行条件变化，导致正常工况波动与故障响应偏离难以区分的问题，本文提出一种基于条件引导扩散模型(Condition-Guided Diffusion Model, CDM)的无监督车辆级故障检测方法。首先，将交叉注意力模块集成至U-Net的低分辨率层和瓶颈层，以充电过程中的电流和荷电状态(State of Charge, SOC)作为条件，引导模型学习给定运行条件下电压和温度的正常响应分布。其次，基于扩散模型的噪声预测误差，构建从片段级异常分数到车辆级异常分数的评估方法，从而在不使用故障样本训练的情况下识别潜在异常。实验结果表明，CDM在Dahu数据集上取得0.8744的AUROC和0.8546的AUPRC，整体优于GP、AE、Deep SVDD、GDN、CATCH、SARAD和DyAD等对比方法。在Dahu、Socea和Naobop三个数据集上，CDM的宏平均AUROC和宏平均AUPRC分别达到0.8806和0.8309，高于主要基准方法DyAD的0.8616和0.8089，表明所提出方法具有较好的车辆级故障综合判别能力。

关键词

锂离子电池，故障检测，扩散模型，条件建模，噪声预测误差

Condition-Guided Diffusion Noise Prediction for Vehicle-Level Fault Detection in Lithium-Ion Batteries

Hong Liu[1,2]

[1]School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan Anhui
[2]Institute of Energy, Hefei Comprehensive National Science Center (Anhui Energy Laboratory), Hefei Anhui

Received: October 5, 2026; accepted: November 4, 2026; published: November 12, 2026

文章引用: 刘红. 基于条件扩散噪声预测的锂离子电池车辆级故障检测[J]. 计算机科学与应用, 2026, 16(11): 117-129.
DOI: 10.12677/csa.2026.1611330

Abstract

The scarcity of fault samples in real-world vehicle charging data, together with the dependence of voltage and temperature responses on operating conditions such as current and state of charge (SOC), makes it difficult to distinguish normal operating fluctuations from fault-induced deviations. To address this issue, this paper proposes an unsupervised vehicle-level fault detection method based on a Condition-Guided Diffusion Model (CDM). First, cross-attention modules are integrated into the low-resolution and bottleneck layers of a U-Net. Charging current and SOC are used as conditions to guide the model in learning the distribution of normal voltage and temperature responses under given operating conditions. Second, an evaluation method is developed based on the noise-prediction error of the diffusion model, aggregating segment-level anomaly scores into vehicle-level anomaly scores and thereby enabling potential anomalies to be identified without using fault samples for training. Experimental results show that CDM achieves an AUROC of 0.8744 and an AUPRC of 0.8546 on the Dahu dataset, outperforming the GP, AE, Deep SVDD, GDN, CATCH, SARAD, and DyAD baselines overall. Across the Dahu, Socea, and Naobop datasets, CDM achieves macro-averaged AUROC and AUPRC values of 0.8806 and 0.8309, respectively, exceeding the corresponding values of 0.8616 and 0.8089 obtained by the primary baseline, DyAD. These results demonstrate that the proposed method provides effective overall discrimination of vehicle-level faults.

Keywords

Lithium-Ion Batteries, Fault Detection, Diffusion Model, Conditional Modeling, Noise Prediction Error

Copyright © 2026 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

随着交通领域电动化和智能化进程不断推进，电动汽车对高性能动力电池的需求持续增长[1]。锂离子电池凭借能量密度高、循环寿命长和自放电率低等特点，已广泛应用于电动汽车及其他低碳能源系统[2]。然而，近年来，全球各地因电池安全问题已召回数十万辆电动汽车[3]。电池在过度充放电、长期运行、高低温环境、挤压、碰撞、针刺或异物穿透等复杂工况下，易发生性能衰退、内短路、热失控等故障[4]。严重时可能引发火灾甚至爆炸[5]。因此，准确的电池故障检测对于预防事故，保障人们的生命财产安全至关重要。

传统的故障检测方法主要包括阈值与统计分析方法、信号处理方法以及模型驱动方法[6]。Xue 等[7]根据运行数据的统计分布构造诊断系数，并结合离群点筛选实现异常单体识别。Li 等[8]采用辛几何模态分解处理电压序列，并结合动态时间规整距离与二维特征聚类识别异常单体。Yi 等[9]利用一阶 RC 等效电路模型和改进扩展卡尔曼滤波器估计 SOC，并依据 SOC 估计偏差检测电池包软短路故障。这些方法虽能在规则明确或模型较为准确的场景下实现有效的早期故障检测，但面对真实车辆中强非线性和多变量耦合特征的复杂电池系统时，仍存在精确模型构建与参数辨识困难、人工特征难以充分表征潜在故障信息，以及诊断规则较为依赖专家经验等局限[10]。

为减少对人工特征和精确机理模型的依赖，深度学习逐渐被应用于工业故障检测。Schaeffer 等[11]利用递归时空高斯过程建模现场运行数据中的时变和工况相关电阻，并构建概率故障检测规则，实现电池

系统的在线健康监测与故障分析。Spina 等[12]比较了多种自编码器(Autoencoder, AE)在工业过程故障检测中的表现，验证了自编码网络在无监督故障检测任务中的适用性。Zhu 等[13]系统研究了变分自编码器(Variational Autoencoder, VAE)及其时序变体在工业过程故障检测与诊断中的应用。Kou 等[14]构建了多尺度深度支持向量数据描述(Deep Support Vector Data Description, Deep SVDD)模型，实现滚动轴承早期故障的在线检测。针对多传感器变量之间的动态耦合关系，Deng 和 Hooi [15]提出图偏差网络(Graph Deviation Network, GDN)，通过图结构学习和图注意力机制刻画变量依赖关系，并利用预测值与观测值之间的偏差检测多变量时间序列异常。

针对锂离子电池故障，研究者进一步提出了多种深度学习故障诊断方法。Yao 等[16]将经验模态分解与卷积神经网络结合，以增强电池包电压故障特征。Sun 等[17]利用 CNN-LSTM 预测电压序列，并根据预测偏差和相关系数识别异常。Liu 等[18]采用 Vision Transformer 和迁移学习实现多类型电池故障诊断。Zhang 等[19]构建了动态自编码器(Dynamical autoencoder for Anomaly Detection, DyAD)框架，通过动态自编码结构和里程弱监督学习车辆充电过程的时序规律，基于重构误差的大小实现了故障车辆的准确识别。Cao 等[20]则将机理约束引入深度网络，以提升随机运行条件下的在线故障诊断能力。

去噪扩散概率模型(Denoising Diffusion Probabilistic Model, DDPM)通过逐步加噪与反向去噪学习复杂数据分布[21]。凭借这一分布建模能力，DDPM 近年来逐渐被用于学习正常数据模式并开展无监督异常检测。Wyatt 等[22]提出 AnoDDPM，通过部分扩散将异常输入映射至正常分布，并利用重构差异定位异常。Zhang 等[23]提出 DiffusionAD，采用一步去噪和范数引导策略提高正常重构质量与推理效率。面向多元时间序列，Xiao 等[24]和 Chen 等[25]分别提出 DiffAD 和 ImDiffusion，利用观测值引导扩散插补并根据去噪结果识别异常。Zhong 等[26]提出 MODEM，通过多分辨率扩散和频域分解建模非平稳时间序列。上述方法取得了良好效果，但锂电池充电时序具有明显的工况依赖性，电压和温度响应受电流与 SOC 共同影响。现有通用方法通常对不同变量进行联合建模，难以显式区分运行条件与系统响应。同时，其异常评分多依赖迭代反向去噪后的重构或插补误差，增加了检测阶段的计算开销。

针对上述特点，本文提出一种基于 CDM 的锂离子电池车辆级故障检测方法。该方法仅使用正常车辆的充电数据训练，将电流和 SOC 作为条件特征，将电压和温度作为响应特征。训练时，条件特征不参与加噪，并经条件编码器和交叉注意力机制注入 U-Net 噪声预测网络。其中，响应特征作为查询，条件特征作为键和值，交叉注意力机制使模型能够根据不同片段和时间位置动态选择相关工况信息。该结构保留了“运行条件–电池响应”的方向性，并为表征电压的快速变化、温度的滞后响应及其随 SOC 变化的非线性关联提供了相应的建模机制。推理时，无需执行完整的反向扩散采样，而是在多个预设扩散时间步上计算注入噪声与预测噪声的均方误差，经平均得到片段级异常分数，再聚合同一车辆的全部片段，得到车辆级异常分数。

本文的主要贡献如下：

(1) 构建仅利用正常车辆数据训练的条件扩散故障检测方法，依据变量的物理功能划分条件特征与响应特征，并通过交叉注意力学习给定电流和 SOC 条件下电压与温度的正常响应规律。

(2) 提出基于多扩散时间步噪声预测误差的异常评分方法，无需完整的反向扩散采样，并将片段级异常分数聚合至车辆级，实现与实际故障标签粒度相匹配的检测。

2. 数据集

2.1. 数据集组成

本文实验使用 Zhang 等[19]公开发表的现实世界电动汽车数据集，该数据集由 Dahu、Socea 和 Naobop 三个子数据集组成，共包含 347 辆电动汽车，约 70 万条充电片段。车辆级故障标签是根据驾驶员的报告

生成的，并由工程师进行确认。车辆数和片段数的数量分布情况如表 1 所示。数据集以汽车编号作为唯一标识，每辆车的数据包含多个充电片段，每个充电片段记录了该车辆在充电过程中的运行状态，并且包含车辆级故障标签。每个充电片段的数据以一个 128 × 7 的矩阵存储，128 代表序列长度，7 表示特征数量。各个特征按表 2 顺序存储。

Table 1. Statistics of vehicles and charging segments in the three datasets
表 1. 三个数据集的车辆与充电片段数量统计

数据集	正常车辆	故障车辆	正常片段	故障片段	总片段
Dahu	168	30	459,708	17,031	476,739
Socea	33	16	82,958	111,287	194,245
Naobop	91	9	27,824	1774	29,598

Table 2. Physical meanings and model roles of charging-segment features
表 2. 充电片段特征的物理含义及其在模型中的角色

序号	特征	含义	本文中的角色
0	Volt	总电压	响应特征
1	Current	电流	条件特征
2	SOC	荷电状态	条件特征
3	Max_single_volt	最高单体电压	响应特征
4	Min_single_volt	最低单体电压	响应特征
5	Max_temp	最高温度	响应特征
6	Min_temp	最低温度	响应特征

2.2. 数据集归一化

本文根据训练集中全部充电片段统计各特征的最小值和最大值，并将训练数据线性映射至 $[-1,1]$。公式如下：

$$x_{\text{norm}} = 2\frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \tag{1}$$

其中，x 是原始特征，$x_{\max}$ 是特征在训练集上的最大值，$x_{\min}$ 是最小值，x_{norm} 则是归一化后的值。测试集使用同一组训练集统计量进行归一化。

3. 条件扩散故障检测方法

3.1. 整体流程

本文提出的条件扩散故障检测与车辆级评估流程如图 1 所示。该流程包括噪声预测与片段级异常评分、车辆级聚合与性能评估两个阶段。首先，将充电数据拆分为响应特征和条件特征，在多个扩散时间步下计算噪声预测误差，并通过多时间步平均获得片段级异常分数；随后，根据车辆编号聚合同一车辆的片段级异常分数，并结合故障标签计算车辆级性能指标。故障标签仅用于性能评价，不参与模型训练和异常分数计算。

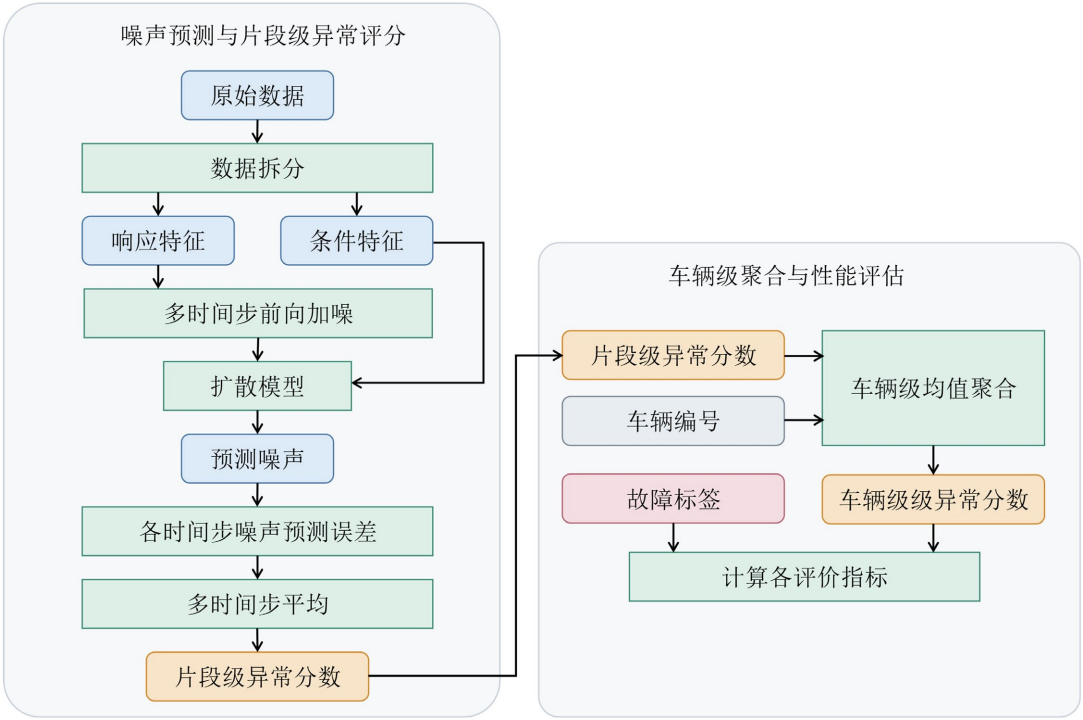


Figure 1. Workflow of conditional diffusion-based fault detection and vehicle-level evaluation
图 1. 条件扩散故障检测与车辆级评估流程图

3.2. 条件扩散过程

本文提出的 CDM 用于锂离子电池故障检测，与通过反向去噪生成或重构样本的方法不同，本文仅利用前向扰动和条件噪声预测进行异常评分，推理阶段不执行迭代反向采样。对于长度为 $L=128$ 的充电片段，将电流和 SOC 作为条件特征：

$$c=\left[I,\mathrm{SOC}\right]\in \mathrm{R}^{L\times 2} \tag{2}$$

将总电压、最高与最低单体电压以及最高与最低温度作为响应特征：

$$x_0=\left[V,V_{\max},V_{\min},T_{\max},T_{\min}\right]\in \mathrm{R}^{L\times 5} \tag{3}$$

训练阶段仅向响应特征加入高斯噪声，条件特征始终保持干净。定义 $\alpha_t=1-\beta_t$ 和 $\bar{\alpha}=\prod_{i=1}^{t}\alpha_i$，任意时间步的噪声响应特征可直接表示为：

$$x_t=\sqrt{\bar{\alpha}_t}x_0+\sqrt{1-\bar{\alpha}_t}\varepsilon,\qquad \varepsilon\sim \mathrm{N}\left(0,\mathrm{I}\right) \tag{4}$$

其中，$t\in\{1,2,\cdots,T\}$ 和 T 分别表示当前扩散时间步与扩散总步数，β_t 为第 t 步的噪声方差，α_t 表示单个扩散时间步的信号保留系数，$\bar{\alpha}_t$ 为从第 1 步到第 t 步的累计信号保留系数，ε 为正向扩散过程中加入的真实标准高斯噪声，I 为单位矩阵。

3.3. 条件噪声预测

噪声预测器采用一维 U-Net，其主干输入为加噪响应特征 x_t 和扩散时间步 t，整体结构如图 2 所示。时间步 t 经过时间嵌入后注入 U-Net 的残差特征，用于表征当前噪声强度。由电流和 SOC 组成的条件特征序列 c 始终保持未加噪状态，并经过条件编码器映射为编码后的条件特征：

$$h_c = f_{\text{cond}}(c) \in \mathrm{R}^{L\times d_c} \tag{5}$$

其中，$f_{\text{cond}}(\cdot)$ 表示条件编码器，h_c 为编码后的条件特征，d_c 表示条件特征维度，本文取 $d_c = 32$。

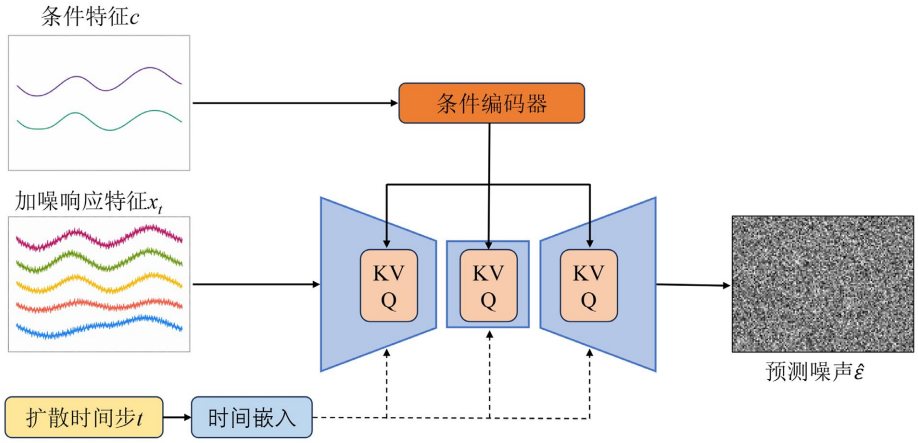


Figure 2. Architecture of the noise prediction network in CDM
图 2. CDM 的噪声预测网络结构

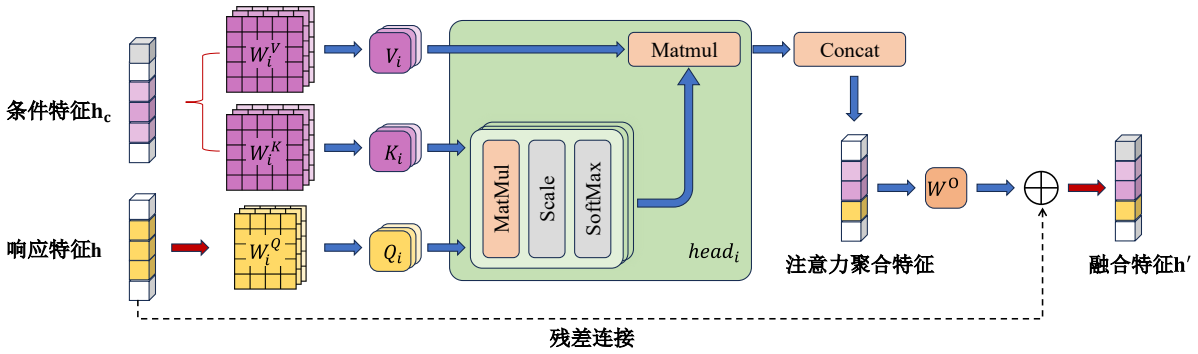


Figure 3. Multi-head cross-attention mechanism guided by condition features
图 3. 条件特征引导的多头交叉注意力机制

为建模运行条件与电压、温度响应之间的动态关联，本文在 U-Net 的低分辨率编码阶段、瓶颈层和相应解码阶段引入多头交叉注意力，其计算过程如图 3 所示。对于第 i 个注意力头，查询、键和值分别为：

$$Q_i = hW_i^{\mathrm{Q}}, \qquad K_i = h_c W_i^{\mathrm{K}}, \qquad V_i = h_c W_i^{\mathrm{V}} \tag{6}$$

其中，h 表示 U-Net 当前层经过时间嵌入和残差计算后的响应特征，W_i^{Q}、W_i^{K} 和 W_i^{V} 为可训练的线性映射参数。交叉注意力输出为：

$$\mathrm{head}_i = \mathrm{softmax}\left(\frac{Q_i K_i^{\mathrm{T}}}{\sqrt{d_h}}\right) V_i \tag{7}$$

$$h' = h + W^{\mathrm{o}}\,\mathrm{Concat}\left(head_1, \cdots, head_H\right) \tag{8}$$

其中，d_h 表示单个注意力头的特征维度，H 表示注意力头数，本文设置 $H = 4$、$d_h = 32$，W^{o} 表示输出映射参数，由核大小为 1 的一维卷积实现，用于将多头拼接结果映射回当前 U-Net 通道维度，h' 表示融

合条件特征后的响应特征。该机制以 U-Net 响应特征作为查询，以编码后的条件特征作为键和值，使模型能够根据电流和 SOC 自适应地提取与当前响应特征相关的条件特征。

经过编码、条件注入和解码后，网络输出对真实噪声的预测：

$$\hat{\varepsilon} = \varepsilon_\theta\left(x_t, t, c\right) \tag{9}$$

其中，$\varepsilon_\theta(\cdot)$ 表示参数为 θ 的条件噪声预测网络，$\hat{\varepsilon}$ 为预测噪声。训练时，从正常车辆数据中随机抽取充电片段，并随机采样扩散时间步 t 和高斯噪声 ε。训练损失为：

$$\mathcal{L} = \left\|\varepsilon - \hat{\varepsilon}\right\|_2^2 \tag{10}$$

3.4. 片段级异常评分

推理阶段在固定的扩散时间步集合 $\mathcal{T}_{\text{eval}} = \{50, 150, 250, \cdots, 950\}$ 上分别执行前向加噪和噪声预测。第 i 个充电片段的异常分数，即该片段的噪声预测误差(Noise Prediction Error, NPE)定义为：

$$s_i = \frac{1}{\left|\mathcal{T}_{\text{eval}}\right|} \sum_{t \in \mathcal{T}_{\text{eval}}} \text{MSE}\left[\varepsilon_{i,t}, \varepsilon_\theta\left(x_{i,t}, t, c_i\right)\right] \tag{11}$$

其中，$\mathcal{T}_{\text{eval}}$ 为推理阶段采用的扩散时间步集合，$\left|\mathcal{T}_{\text{eval}}\right|$ 为其元素个数，MSE 为均方误差函数。多个扩散时间步的误差取平均，可以综合不同噪声强度下的预测结果，并减小单一时间步和单次随机噪声对异常分数的影响。片段级异常分数 s_i 越大，说明该片段越偏离模型学习到的条件正常分布，其异常程度越高。

3.5. 车辆级异常评分

推理阶段首先根据上式计算各充电片段的异常分数。对于第 j 辆车，将其全部充电片段的异常分数取平均，得到车辆级异常分数：

$$S_j = \frac{1}{\left|\mathcal{J}_j\right|} \sum_{i \in \mathcal{J}_j} s_i \tag{12}$$

其中，$\mathcal{J}_j$ 表示第 j 辆车所包含的充电片段索引集合，$\left|\mathcal{J}_j\right|$ 表示该车辆的充电片段数量。S_j 越大，表明该车辆的电压和温度响应偏离条件正常分布的程度越高。

4. 实验与结果分析

4.1. 实验环境

本文所有实验基于 Python3.8 和 PyTorch 框架，使用 NVIDIA GeForce RTX4090 的 GPU 和 24 GB 的显存。扩散过程的时间步设置为 $T = 1000$，定义正向扩散过程加入的噪声方差 $\beta_1 = 0.0001$，$\beta_T = 0.02$，噪声强度随时间步线性增加。训练过程中使用的优化器是 AdamW，权重衰减系数设置为 0.0001。设置初始学习率和训练轮次分别为 0.0001 和 50，学习率预热倍数为 5，选取前 10%的 epoch 用于预热，预热结束后采用余弦退火方法减小学习率。用于训练的批量大小设置为 128，推理过程的批量大小同样设置为 128。

为保证评价的可靠性，本文采用车辆级五折交叉验证，正常车辆和故障车辆分别划分为 5 折，每次使用 4 折正常车辆训练，为了平衡测试集中正常车辆与故障车辆数量，剩余 1 折正常车辆与当前折之外 4 折故障车辆共同构成测试集。为降低随机初始化和训练随机性对结果的影响，设置 5 组随机种子，每组均完成五折实验。实验采用 3、13、23、33、43 五组训练随机种子，推理过程分别使用 1003、1013、1023、1033、1043 作为对应的评价种子。

4.2. 评价指标

为全面评价车辆级故障检测性能，本文采用受试者工作特征曲线下面积(Area Under the Receiver Operating Characteristic Curve, AUROC)和精确率–召回率曲线下面积(Area Under the Precision-Recall Curve, AUPRC)作为主要评价指标，并报告精确率(Precision)、召回率(Recall)和 F1 分数。其中，故障车辆定义为正类，正常车辆定义为负类。AUROC 用于衡量模型在不同阈值下区分正常车辆与故障车辆的整体能力，AUPRC 进一步反映类别不均衡情况下对故障车辆的识别性能。宏平均 AUROC 和 AUPRC 为三个数据集五折均值的算术平均，各数据集等权。计算 Precision、Recall 和 F1 时，以各方法在正常训练车辆上所得车辆级异常分数的第 95 百分位数作为判定阈值。上述指标的值越大，表明模型检测性能越好。

4.3. 对比实验

4.3.1. Dahu 主要基准结果

Table 3. Vehicle-level fault detection results of different methods on the Dahu dataset
表 3. 不同方法在 Dahu 数据集上的车辆级故障检测结果

方法	AUROC	AUPRC	Precision (%)	Recall (%)	F1 (%)
CATCH [27]	0.5024 ± 0.0815	0.4803 ± 0.0694	57.00 ± 16.09	10.00 ± 2.28	16.91 ± 3.83
GP	0.5179 ± 0.0743	0.4863 ± 0.0619	58.33 ± 13.74	10.00 ± 2.28	17.00 ± 3.73
SARAD [28]	0.5199 ± 0.0724	0.4989 ± 0.0484	57.90 ± 15.14	13.33 ± 3.49	21.58 ± 5.54
Deep SVDD	0.5212 ± 0.1229	0.4529 ± 0.0947	40.00 ± 25.28	5.83 ± 4.75	9.98 ± 7.67
GDN	0.6303 ± 0.1374	0.6242 ± 0.1550	61.47 ± 38.91	25.00 ± 19.76	34.80 ± 26.22
AE	0.7763 ± 0.0338	0.7470 ± 0.0561	80.64 ± 14.89	30.00 ± 6.18	43.50 ± 7.84
DyAD	0.8432 ± 0.0501	0.8470 ± 0.0617	**86.34** ± 8.74	**60.83** ± 7.57	**71.13** ± 6.70
CDM(Ours)	**0.8744** ± 0.0416	**0.8546** ± 0.0664	84.03 ± 14.61	60.00 ± 6.97	69.02 ± 4.16

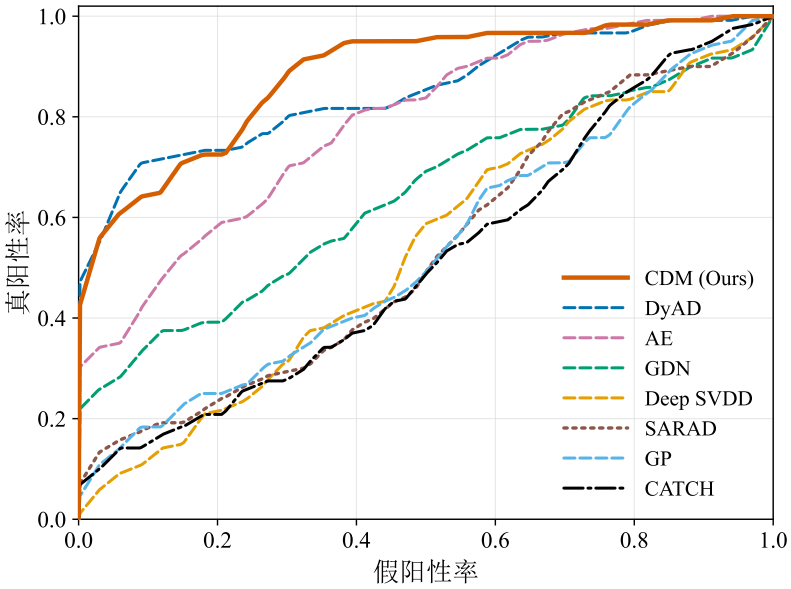


Figure 4. Five-fold mean ROC curves of eight methods on the Dahu dataset
图 4. Dahu 数据集上八种方法的五折平均 ROC 曲线

为验证 CDM 的检测效果，将所提出的方法与 7 种对比方法在 Dahu 数据集上进行比较。为保证评价的一致性，所有方法采用相同的车辆级五折数据划分，训练随机种子设为 3，结果如表 3 所示。表中加粗数据表示最优结果，下划线数据表示次优结果。

从对比结果来看，CDM 取得了 0.8744 的平均 AUROC 和 0.8546 的平均 AUPRC，均为所有方法中的最优结果，与性能次优的 DyAD 相比，分别提高了 0.0312 和 0.0076，即 3.12 和 0.76 个百分点。在固定阈值下，DyAD 在精确率、召回率和 F1 分数上均取得最优结果，CDM 在三项指标上均位列第二，表明 CDM 具有较强的整体异常排序能力，并在给定阈值下保持了较好的分类性能。为更加直观地比较各方法的检测效果，图 4 进一步给出了相应的五折平均 ROC 曲线。CDM 的平均 ROC 曲线整体更接近左上角，曲线下面积最大，与表 3 中的定量结果一致。AE 表现出一定的车辆级判别能力，但整体低于 CDM 和 DyAD，而 GP 和 Deep SVDD 的 AUROC 接近 0.5，在该任务上的车辆级判别能力有限，GDN 虽有所改善，但表现出较大的折间波动。非扩散方法 CATCH 和 SARAD 的 AUROC 接近 0.5，可能与这两种方法适用于异常发生位置较为明确的点级或子序列级时序异常检测，而本文数据集仅提供车辆级故障标签，缺少故障发生时刻以及点级或片段级异常标注有关。

4.3.2. Socea 和 Naobop 数据集验证

Table 4. Comparison of vehicle-level AUROC and macro-average results of CDM and DyAD across three datasets
表 4. CDM 与 DyAD 在三个数据集上的车辆级 AUROC 及宏平均比较

方法	Dahu	Socea	Naobop	Macro-AUROC
DyAD	0.8460 ± 0.0539	**0.9100** ± 0.0160	0.8289 ± 0.0603	0.8616
CDM	**0.8757** ± 0.0392	0.9093 ± 0.0383	**0.8567** ± 0.0730	**0.8806**

Table 5. Comparison of vehicle-level AUPRC and macro-average results of CDM and DyAD across three datasets
表 5. CDM 与 DyAD 在三个数据集上的车辆级 AUPRC 及宏平均比较

方法	Dahu	Socea	Naobop	Macro-AUPRC
DyAD	0.8464 ± 0.0615	**0.9549** ± 0.0167	0.6253 ± 0.0738	0.8089
CDM	**0.8580** ± 0.0567	0.9400 ± 0.0488	**0.6945** ± 0.1028	**0.8309**

为进一步比较 CDM 与基准方法 DyAD 的检测性能，在 Dahu、Socea 和 Naobop 三个数据集上对车辆级 AUROC 和 AUPRC 进行评估，结果如表 4 和表 5 所示。表中结果来自 5 组随机种子下的五折实验，共包含 25 个种子–折次结果。CDM 在 Dahu 和 Naobop 上的 AUROC 与 AUPRC 均高于 DyAD，在 Socea 上，两种方法的 AUROC 基本持平，而 DyAD 的 AUPRC 略高。综合三个数据集，CDM 的宏平均 AUROC 和宏平均 AUPRC 分别达到 0.8806 和 0.8309，较 DyAD 的 0.8616 和 0.8089 分别提高 1.90 和 2.20 个百分点。上述结果说明，CDM 性能基本优于基准方法 DyAD。

4.4. 消融实验

为分析条件信息和交叉注意力的作用，设置三个消融版本。无条件扩散：移除条件编码器和全部交叉注意力块，仅根据 $\boldsymbol{x}_t$ 和 t 预测噪声；条件盲化： 保留完整条件模型和参数规模，但在训练、验证和推理时将电流和 SOC 统一替换为 0；条件拼接：将未加噪的电流和 SOC 作为条件直接与加噪特征 $\boldsymbol{x}_t$ 输入网络，消融实验在三个数据集上使用与主模型一致的 5 个种子和 5 个折次，实验结果如表 6 和表 7 所示。

从表 6 可以看出，无条件扩散与条件盲化在三个数据集上的 AUROC 较为接近，说明仅保留条件建模结构而不提供真实条件信息，难以带来明显的性能提升。条件拼接在三个数据集上的 AUROC 均高于

上述两种消融版本，表明电流和 SOC 能够为异常检测提供有效的运行条件信息。完整 CDM 在三个数据集上均取得最高 AUROC，其宏平均 AUROC 达到 0.8806，较无条件扩散、条件盲化和条件拼接分别提高 10.82、10.41 和 4.98 个百分点。表 7 中的 AUPRC 结果呈现出相同趋势，CDM 的宏平均 AUPRC 达到 0.8309，较三种消融版本分别提高 8.86、8.33 和 5.12 个百分点。上述结果表明，模型性能的提升不仅来源于电流和 SOC 所提供的真实运行条件，还与条件信息的注入方式有关，相比直接拼接，交叉注意力能够更充分地建模运行条件与电池响应之间的动态关联，从而提升车辆级故障检测性能。

Table 6. AUROC results of the conditioning-information ablation study
表 6. 条件信息消融实验的 AUROC 结果

方法	Dahu	Socea	Naobop	Macro-AUROC
无条件扩散	0.7451 ± 0.0760	0.7434 ± 0.1085	0.8288 ± 0.0731	0.7724
条件盲化	0.7475 ± 0.0762	0.7439 ± 0.1117	0.8380 ± 0.0735	0.7765
条件拼接	0.8003 ± 0.0366	0.8412 ± 0.0383	0.8508 ± 0.0699	0.8308
CDM	**0.8757** ± 0.0392	**0.9093** ± 0.0383	**0.8567** ± 0.0730	**0.8806**

Table 7. AUPRC results of the conditioning-information ablation study
表 7. 条件信息消融实验的 AUPRC 结果

方法	Dahu	Socea	Naobop	Macro-AUPRC
无条件扩散	0.7271 ± 0.0791	0.8475 ± 0.0910	0.6523 ± 0.0913	0.7423
条件盲化	0.7306 ± 0.0785	0.8480 ± 0.0907	0.6641 ± 0.0961	0.7476
条件拼接	0.7898 ± 0.0541	0.8606 ± 0.0483	0.6887 ± 0.0992	0.7797
CDM	**0.8580** ± 0.0567	**0.9400** ± 0.0488	**0.6945** ± 0.1028	**0.8309**

4.5. 噪声预测误差分布与可分性分析

为直观分析 CDM 噪声预测误差的故障表征能力，本文以 Dahu 数据集随机种子 3、第 0 折的结果为例，对训练集正常样本、测试集正常样本和测试集故障样本的 NPE 分布进行可视化，如图 5 所示。

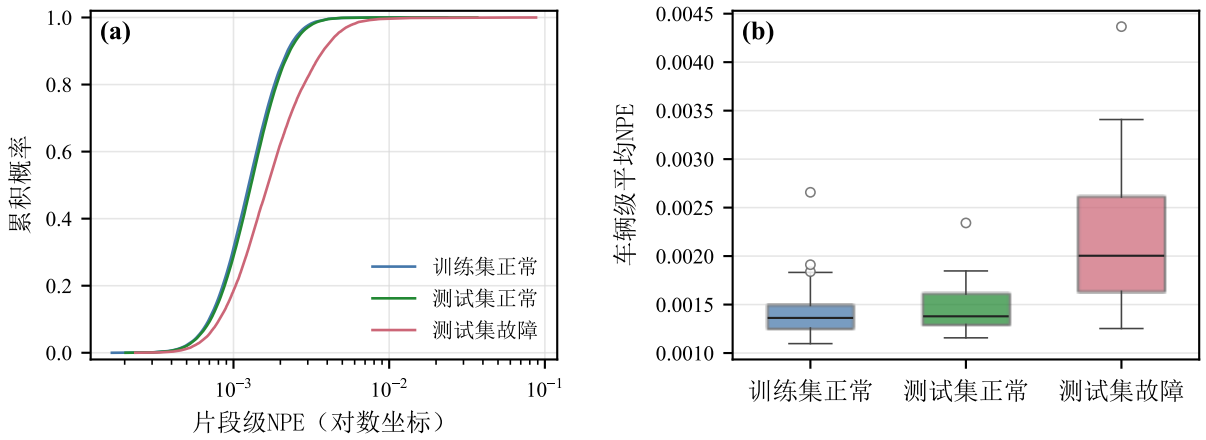


Figure 5. Distribution of noise prediction errors on the Dahu dataset: (a) empirical cumulative distribution of segment-level NPE; (b) box plot of vehicle-level mean NPE
图 5. Dahu 数据集上的噪声预测误差分布：(a) 片段级 NPE 经验累积分布；(b) 车辆级平均 NPE 箱线图

图 5(a)给出了片段级 NPE 的经验累积分布。训练集正常样本与测试集正常样本的曲线基本重合，表明模型在未参与训练的正常车辆上产生了与训练数据相近的噪声预测误差分布。相比之下，测试集故障样本的曲线整体向高 NPE 方向偏移，且较大预测误差片段出现得更加频繁。这说明当电压和温度响应偏离给定电流和 SOC 条件下的正常变化规律时，模型的噪声预测误差通常会增大。

图 5(b)进一步展示了车辆级平均 NPE 的分布。训练集正常车辆与测试集正常车辆的分布较为接近，而测试集故障车辆的 NPE 整体较高，并与正常车辆呈现出较明显的分离趋势。对同一车辆的片段级 NPE 进行平均，可以削弱单个片段随机波动的影响，并累积车辆中持续出现的异常响应，有助于增强车辆级故障的可分性。

需要指出的是，正常车辆与故障车辆的 NPE 分布仍存在一定重叠。这是因为故障标签属于车辆级标签，故障车辆并非所有充电片段都表现出明显异常。因此，本文不使用 NPE 进行片段级硬分类，而是将片段级 NPE 聚合为车辆级异常分数。图 5 的分布结果直观说明了采用 NPE 表征异常并进行车辆级聚合的合理性。

4.6. 成本分析

Table 8. Comparison of model complexity and computational costs across different methods
表 8. 不同方法的模型复杂度与计算成本比较

方法	参数量(M)	模型大小(MiB)	训练时间(min/折)	推理延迟(ms/片段)
DyAD	0.8	3.1	1.9 ± 0.1	1.4 ± 0.2
CATCH [27]	26.2	100.0	18.2 ± 1.0	5.2 ± 0.2
SARAD [28]	9.5	36.4	18.2 ± 6.6	1.3 ± 0.1
CDM (Ours)	2.2	8.5	128.6 ± 6.9	110.5 ± 2.9

为了综合评估不同种类方法的计算与存储开销，本文进一步比较几种方法的参数量、模型大小、训练时间和推理延迟，结果如表 8 所示。CDM 的参数量为 2.2 M，模型大小为 8.5 MiB，虽高于 DyAD，但明显低于 CATCH 和 SARAD，整体保持了较为紧凑的模型规模。CDM 的训练时间和推理延迟相对较高，主要是因为需要在多个预设扩散时间步上重复进行噪声预测，但同时在承担额外计算开销的同时，CDM 取得了更高的 AUROC 和 AUPRC，体现出较好的检测性能与模型规模平衡。考虑到模型训练在离线阶段完成，且本文面向车辆充电数据的服务器端批量检测，当前计算开销具有一定的应用可行性。后续可通过减少评估时间步、模型压缩和混合精度推理进一步提升计算效率。

4.7. 未来研究

本文方法仍依赖于条件特征与响应特征的先验划分，其适用性可能随车型和运行工况发生变化。受限于车辆级标签，当前尚难以评估模型对不同故障类型、阶段和严重程度的敏感性，同时，均值聚合可能稀释少量异常片段的贡献。未来将探索具有物理约束的端到端特征划分和基于注意力或多实例学习的车辆级聚合方法，并结合更细粒度的故障标签提升检测与定位能力。此外，多扩散时间步上的重复噪声预测增加了模型的训练和推理开销。未来将研究自适应时间步选择、少时间步评分及模型轻量化方法，以提升其在资源受限场景下的部署效率。

5. 结论

本文提出了一种面向真实车辆充电数据的条件扩散故障检测方法。该方法将电流和 SOC 作为运行条件，仅对电压和温度响应特征进行扩散，并利用一维 U-Net 预测不同噪声尺度下的响应噪声。推理阶段

先对多个扩散时间步的噪声预测误差取平均，得到片段级异常分数，再聚合同一车辆的片段级异常分数形成车辆级异常分数。固定随机种子下的五折对比实验表明，CDM 在 Dahu 数据集上取得 0.8744 ± 0.0416 的车辆级 AUROC 和 0.8546 ± 0.0664 的车辆级 AUPRC，优于 7 种基准方法。在 5 组随机种子、每组 5 折的重复实验中，CDM 在三个数据集上的宏平均 AUROC 和宏平均 AUPRC 分别为 0.8806 和 0.8309，高于 DyAD 的 0.8616 和 0.8089。消融实验表明，交叉注意力能够根据不同时间位置和响应状态自适应地提取相关工况信息，从而更充分地建模电流和 SOC 与电压、温度响应之间的动态关联，提升车辆级故障检测性能。总体而言，条件扩散噪声预测为区分正常工况变化与故障响应偏离提供了一种可行方法。

基金项目

国家自然科学基金资助项目(52503291)；合肥综合性国家科学中心能源研究院(安徽省能源实验室)启动基金(24KZS501)。

参考文献

[1] Shang, Y., Wang, S., Tang, N., Fu, Y. and Wang, K. (2024) Research Progress in Fault Detection of Battery Systems: A Review. *Journal of Energy Storage*, **98**, Article 113079. https://doi.org/10.1016/j.est.2024.113079

[2] Zhao, J., Liu, M., Zhang, B., Wang, X., Liu, D., Wang, J., *et al*. (2024) Review of Lithium-Ion Battery Fault Features, Diagnosis Methods, and Diagnosis Procedures. *IEEE Internet of Things Journal*, **11**, 18936-18950. https://doi.org/10.1109/jiot.2023.3324322

[3] Zhao, J., Feng, X., Wang, J., Lian, Y., Ouyang, M. and Burke, A.F. (2023) Battery Fault Diagnosis and Failure Prognosis for Electric Vehicles Using Spatio-Temporal Transformer Networks. *Applied Energy*, **352**, Article 121949. https://doi.org/10.1016/j.apenergy.2023.121949

[4] Kong, D., Lv, H., Ping, P. and Wang, G. (2023) A Review of Early Warning Methods of Thermal Runaway of Lithium Ion Batteries. *Journal of Energy Storage*, **64**, Article 107073. https://doi.org/10.1016/j.est.2023.107073

[5] Yu, Q., Wang, C., Li, J., Xiong, R. and Pecht, M. (2023) Challenges and Outlook for Lithium-Ion Battery Fault Diagnosis Methods from the Laboratory to Real World Applications. *eTransportation*, **17**, Article 100254. https://doi.org/10.1016/j.etran.2023.100254

[6] Zou, B., Zhang, L., Xue, X., Tan, R., Jiang, P., Ma, B., *et al*. (2023) A Review on the Fault and Defect Diagnosis of Lithium-Ion Battery for Electric Vehicles. *Energies*, **16**, Article 5507. https://doi.org/10.3390/en16145507

[7] Xue, Q., Li, G., Zhang, Y., Shen, S., Chen, Z. and Liu, Y. (2021) Fault Diagnosis and Abnormality Detection of Lithium-Ion Battery Packs Based on Statistical Distribution. *Journal of Power Sources*, **482**, Article 228964. https://doi.org/10.1016/j.jpowsour.2020.228964

[8] Li, S., Zhang, C., Du, J., Cong, X., Zhang, L., Jiang, Y., *et al*. (2022) Fault Diagnosis for Lithium-Ion Batteries in Electric Vehicles Based on Signal Decomposition and Two-Dimensional Feature Clustering. *Green Energy and Intelligent Transportation*, **1**, Article 100009. https://doi.org/10.1016/j.geits.2022.100009

[9] Yi, L., Cai, X., Wang, Y., Luo, B., Liu, J. and Liu, B. (2023) Soft Short-Circuit Fault Diagnosis of the Lithium-Ion Battery Pack Based on an Improved EKF Algorithm. *Recent Patents on Mechanical Engineering*, **16**, 138-149. https://doi.org/10.2174/2212797616666230330103334

[10] Zhang, Y., Jiang, M., Zhou, Y., Zhao, S. and Yuan, Y. (2023) Towards High-Safety Lithium-Ion Battery Diagnosis Methods. *Batteries*, **9**, Article 63. https://doi.org/10.3390/batteries9010063

[11] Schaeffer, J., Lenz, E., Gulla, D., Bazant, M.Z., Braatz, R.D. and Findeisen, R. (2024) Gaussian Process-Based Online Health Monitoring and Fault Analysis of Lithium-Ion Battery Systems from Field Data. *Cell Reports Physical Science*, **5**, Article 102258. https://doi.org/10.1016/j.xcrp.2024.102258

[12] Spina, D.E., de O. Campos, L.F., de Arruda, W.F., Melo, A., Alves, M.F.D.S., Rabello, G.L., *et al*. (2024) Comparison of Autoencoder Architectures for Fault Detection in Industrial Processes. *Digital Chemical Engineering*, **12**, Article 100162. https://doi.org/10.1016/j.dche.2024.100162

[13] Zhu, J., Jiang, M. and Liu, Z. (2022) Fault Detection and Diagnosis in Industrial Processes with Variational Autoencoder: A Comprehensive Study. *Sensors*, **22**, Article 227. https://doi.org/10.3390/s22010227

[14] Kou, L., Chen, J., Qin, Y. and Mao, W. (2022) The Robust Multi-Scale Deep-SVDD Model for Anomaly Online Detection of Rolling Bearings. *Sensors*, **22**, Article 5681. https://doi.org/10.3390/s22155681

[15] Deng, A. and Hooi, B. (2021) Graph Neural Network-Based Anomaly Detection in Multivariate Time Series. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 4027-4035. https://doi.org/10.1609/aaai.v35i5.16523

[16] Yao, L., Zheng, J., Xiao, Y., Zhang, C., Zhang, L., Gong, X., *et al*. (2023) An Intelligent Fault Diagnosis Method for Lithium-Ion Battery Pack Based on Empirical Mode Decomposition and Convolutional Neural Network. *Journal of Energy Storage*, **72**, Article 108181. https://doi.org/10.1016/j.est.2023.108181

[17] Sun, J., Ren, S., Shang, Y., Zhang, X., Liu, Y. and Wang, D. (2023) A Novel Fault Prediction Method Based on Convolutional Neural Network and Long Short-Term Memory with Correlation Coefficient for Lithium-Ion Battery. *Journal of Energy Storage*, **62**, Article 106811. https://doi.org/10.1016/j.est.2023.106811

[18] Liu, X., Cai, H., Zhou, Z., Kong, Y., Zhou, X., Han, X., *et al*. (2024) Enhancing Multi-Type Fault Diagnosis in Lithium-Ion Battery Systems: Vision Transformer-Based Transfer Learning Approach. *Journal of Power Sources*, **624**, Article 235610. https://doi.org/10.1016/j.jpowsour.2024.235610

[19] Zhang, J., Wang, Y., Jiang, B., He, H., Huang, S., Wang, C., *et al*. (2023) Realistic Fault Detection of Li-Ion Battery via Dynamical Deep Learning. *Nature Communications*, **14**, Article No. 5940. https://doi.org/10.1038/s41467-023-41226-5

[20] Cao, R., Zhang, Z., Shi, R., Lu, J., Zheng, Y., Sun, Y., *et al*. (2025) Model-Constrained Deep Learning for Online Fault Diagnosis in Li-Ion Batteries over Stochastic Conditions. *Nature Communications*, **16**, Article No. 1651. https://doi.org/10.1038/s41467-025-56832-8

[21] Ho, J., Jain, A. and Abbeel, P. (2020) Denoising Diffusion Probabilistic Models. 2020 *Advances in Neural Information Processing Systems*, Online, 6-12 December 2020, 6840-6851. https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html

[22] Wyatt, J., Leach, A., Schmon, S.M. and Willcocks, C.G. (2022) AnoDDPM: Anomaly Detection with Denoising Diffusion Probabilistic Models Using Simplex Noise. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (*CVPRW*), New Orleans, 19-20 June 2022, 649-655. https://doi.org/10.1109/cvprw56347.2022.00080

[23] Zhang, H., Wang, Z., Zeng, D., Wu, Z. and Jiang, Y.G. (2025) DiffusionAD: Norm-Guided One-Step Denoising Diffusion for Anomaly Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **47**, 7140-7152. https://doi.org/10.1109/tpami.2025.3570494

[24] Xiao, C., Gou, Z., Tai, W., Zhang, K. and Zhou, F. (2023) Imputation-Based Time-Series Anomaly Detection with Conditional Weight-Incremental Diffusion Models. *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Long Beach, 6-10 August 2023, 2742-2751. https://doi.org/10.1145/3580305.3599391

[25] Chen, Y., Zhang, C., Ma, M., Liu, Y., Ding, R., Li, B., *et al*. (2023) ImDiffusion: Imputed Diffusion Models for Multivariate Time Series Anomaly Detection. *Proceedings of the VLDB Endowment*, **17**, 359-372. https://doi.org/10.14778/3632093.3632101

[26] Zhong, G., Wang, P., Yuan, J., Li, Z. and Chen, L. (2025) Multi-Resolution Decomposable Diffusion Model for Non-Stationary Time Series Anomaly Detection. *International Conference on Learning Representations*, Singapore, 24-28 April 2025, 33527-33553. https://proceedings.iclr.cc/paper_files/paper/2025/hash/9f7f063144103bf6debb09a3f15e00fb-Abstract-Conference.html

[27] Wu, X., Qiu, X., Li, Z., Wang, Y., Hu, J., Guo, C., *et al*. (2025) CATCH: Channel-Aware Multivariate Time Series Anomaly Detection via Frequency Patching. *International Conference on Learning Representations*, Vancouver, 24-28 April 2025, 17017-17045. https://proceedings.iclr.cc/paper_files/paper/2025/hash/2b25c39788e5cf11d3541de433ebf4c0-Abstract-Conference.html

[28] Dai, Z., He, L., Yang, S.H. and Leeke, M. (2024) SARAD: Spatial Association-Aware Anomaly Detection and Diagnosis for Multivariate Time Series. *Advances in Neural Information Processing Systems* 37, Vancouver, 10-15 December 2024, 48371-48410. https://doi.org/10.52202/079017-1533