

Anomaly detection of power battery pack using gated recurrent units based variational autoencoder

Changcheng Sun, Zhiwei He, Huipin Lin, Linhui Cai, Hui Cai, Mingyu Gao

College of electronic information, Hangzhou Dianzi University, Hangzhou 310018, PR China

School of Electrical Engineering, Yancheng Institute of Technology, Yancheng 224051, PR China

目录 CONTENTS

一、摘要

二、背景

三、算法

四、评估

五、结论

六、

一、摘要

1、摘要

快速准确地检测电池组异常并及时进行容错控制对于电动汽车（EV）的安全运行具有重要意义。

提出了一种半监督异常检测模型，通过使用基于门控循环单元（GRU）的变分自动编码器（VAE）（GRU-VAE）框架来检测电动汽车动力电池组的早期潜在异常。

GRU 用于捕获 MVTs 中的复杂时间依赖性，VAE 用于以概率重建输入样本。采用经典极值理论（EVT）中的峰值超过阈值（POT）模型来正确设置异常检测阈值。

在真实电动汽车运行数据集上的实验表明，该模型能够有效检测 MVTs 中潜在的异常情况，有望为智能动力电池组异常检测技术提供参考。

二、背景

1、问题

锂离子电池(LIB) 在电动汽车领域得到广泛应用。然而，锂离子电池的抗滥用能力差，易受外部环境影响，容易出现性能下降、老化，甚至发生燃烧、爆炸等灾难性事故。一辆电动汽车动力电池组是由成百上千个锂离子电池电芯串、并联而成。长期运行会导致电池之间的性能不一致，可能会导致各种严重事故危害。

电池组和整车的监测信号需要集成用以反映电池组的整体状况。只有考虑多维信号以及它们之间的相关性，才能检测到一些潜在的异常行为。

在电池组异常检测的实践中，收集标记为各种异常的数据是困难且昂贵的，而正常运行数据是海量的。

由于外部环境和驾驶习惯的影响，锂离子电池及其系统的行为和规律很难预测。

2、其他人的解决方案

提出了一种结合长短记忆递归神经网络（LSTM）和等效电路模型（ECM）的异常电池诊断模型，可以准确定位潜在的热失控电池。

利用机器学习算法和 3σ 多级筛选策略（ 3σ -MSS）来检测电池组中电芯端电压的异常变化，并以概率的形式进行计算。

提出了一种多尺度卷积循环编码器-解码器（MSCRED），用于在多元时间序列数据中执行异常检测和诊断。

三、算法

1、解决方法

采用半监督深度重构模型GRU-VAE来检测动力电池组的早期异常。GRU能够捕获输入时间序列中复杂的时间依赖性，而VAE则对多维输入序列进行概率重构。并使用POT模型正确设置阈值来识别异常事件。

这项工作为动力电池组提供了一种通用的异常检测框架，对于保证电池组的安全运行将起到关键作用。

数据：国家电动汽车服务管理中心（NSMC-EV）的真实电动汽车运行大数据。

电池管理系统（BMS）利用部署的传感器网络实时测量电池芯的电压、电流、温度等多项物理参数，并将其与车辆行驶数据（如车速、里程、状态等）相结合，形成MVTs，为电动汽车动力电池系统中的技术异常检测提供有价值的数据集。

2、GRU

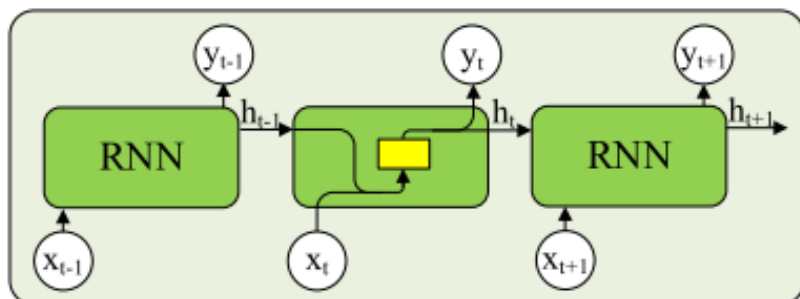
基于循环神经网络（RNN）的模型具有良好的时间序列特征提取能力。

LSTM作为一种改进的RNN，是为了解决传统RNN对长期输入信息不敏感以及反向传播算法中的梯度消失和梯度爆炸问题而提出的，并已广泛应用于UVTS和MVTs的处理中。

LSTM在RNN的基础上增加了门控机制，包括输入门 i_t 、遗忘门 f_t 和输出门 o_t ，分别用于更新、保留和删除循环单元中的信息。

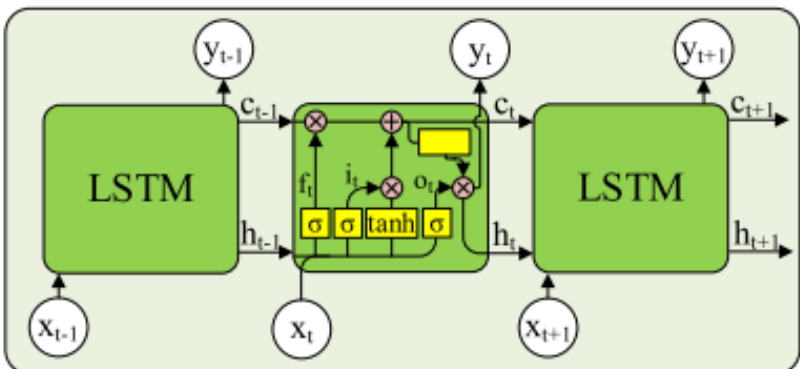
GRU 是 LSTM 的一种变体，结构上比 LSTM 少一个输入控制门，它保留了更新门 u_t 和重置门 r_t 来控制前一个时间戳的状态信息转换为当前时间戳的程度和确定最后隐藏状态中有多少信息需要被遗忘。

2、GRU



$$h_t = \sigma(Wx_t + Uh_{t-1} + b_h)$$

$$y_t = \sigma(Vh_t)$$



$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$

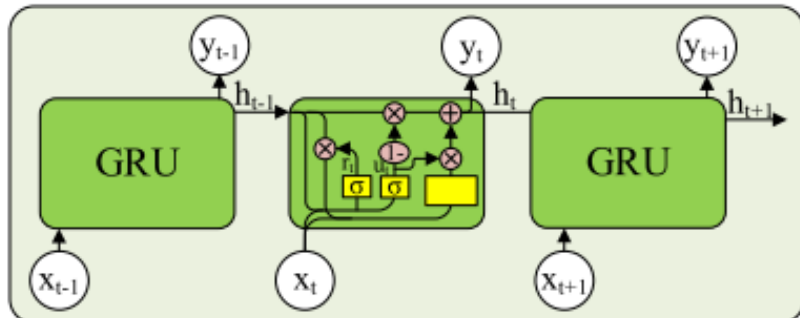
$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$

$$c_t^* = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes c_t^*$$

$$h_t = o_t \otimes \tanh(c_t)$$



$$u_t = \sigma(W_u x_t + U_u h_{t-1} + b_u)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r)$$

$$h_t^* = \tanh(W_h x_t + U_h (r_t \otimes h_{t-1}) + b_h)$$

$$h_t = (1 - u_t) \otimes h_{t-1} + u_t \otimes h_t^*$$

RNN、LSTM 和 GRU 的单元结构以及每个门在时间步 t 的记忆更新公式的图示。 x_t 表示为输入， y_t 表示为输出， h_t 表示隐藏状态， c_t 表示记忆单元。 σ 表示非线性激活函数； W 、 U 和 V 是所有时间步共享的权重矩阵； b 表示不同权重矩阵对应的偏置向量。 \otimes 表示逐元素乘积。

$(1 - u_t) \otimes h_{t-1}$ 表示对原始隐藏状态的选择性“遗忘”，而 $u_t \otimes h_t^*$ 表示包含 h_t^* 的选择性“记住”当前节点的信息。GRU 的参数比标准 LSTM 少，但其泛化性能与 LSTM 相当，有利于提高训练效率。

3、基于GRU的VAE框架

将 ϕ 和 θ 分别表示编码器和解码器的参数集

编码器将具有多个时间步长的多元观测值 x 映射到服从某种概率分布（例如高斯分布）的潜在空间，潜在变量 z 存储为后验分布 $p_{\theta}(z|x)$ 。

通过使用 KL散度，将 $p_{\theta}(z|x)$ 限制为尽可能逼近正态分布 $N(\mu, \sigma^2)$ ，并通过分别使用 ReLU 和 SoftPlus 激活函数的两个并行密集层从 GRU 输出层学习。然后使用重新参数化技术从确定的后验正态分布 $N(z|\mu, \sigma^2 I)$ 中随机采样 z ：

$$z = g_{\theta}(x, \varepsilon) = \mu + \sigma \cdot \varepsilon$$

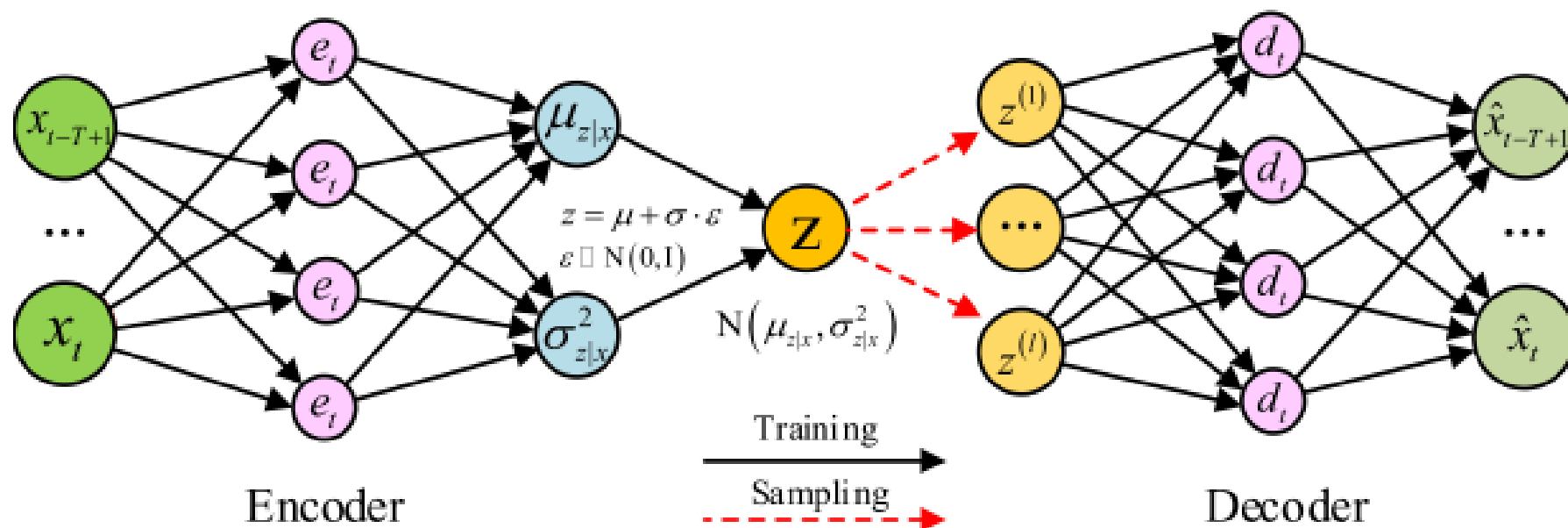
其中 ε 是辅助噪声变量， $\varepsilon \sim N(0, 1)$ 。所有 z 形成一个 Z 空间。

3、基于GRU的VAE框架

解码器从潜在分布中随机采样元素，并根据条件概率分布 $q_\phi(\hat{x}|z)$ 的训练参数生成与 x 相似的重构数据 \hat{x} 。基于随机梯度变分贝叶斯 (SGVB) 理论，VAE的损失函数定义如下：

$$\mathcal{L} = \max_{\theta, \phi} \left\{ \mathbb{E}_{z \sim p_\theta(z|x)} [\ln q_\phi(\hat{x}|z)] - D_{\text{KL}}[p_\theta(z|x) \parallel p_\theta(z)] \right\}$$

第一项为重构概率，用于评价 \hat{x} 与 x 之间的相似度，第二项为正则化项，描述先验分布 $p_\theta(z)$ 与后验分布 $p_\theta(z|x)$ 之间的相似度。



4、算法

使用最小-最大归一化方法对每个MVTs进行归一化。应用滑动时间窗技术将MVTs划分为窗口长度（WL）为T的重叠观测窗口，用于模型的输入更新。相邻窗口允许重叠，间接增加了训练样本量，有助于增强模型的泛化能力。通过调整窗口移动步长（WSS）来控制重叠程度。

Algorithm 1. GRU-based VAE training algorithm

Input: Anomaly-free MVTs: $\mathcal{X}_{train} = \{\omega_T, \omega_{T+b}, \dots, \omega_{T+L \cdot b}\} (\in \mathbb{R}^{N \times M})$

Output: $\phi, \theta \leftarrow$ parameters of Encoder and Decoder

Step:

for $i=0$ to L do:

Update the latent variable of encoder: $h_t^e = (1 - c_t) \otimes h_{t-1}^e + c_t \otimes \tanh(W_h \cdot \mathcal{X}_{train}^{(i)} + U_h \cdot (r_t \cdot h_{t-1}^e) + b_h^e)$;

Encoder calculates the parameters of the posterior distribution of latent variable z :

$$\begin{cases} \mu_{z_t} = W_{\mu_z} h_t^e + b_{\mu_z} \\ \ln \sigma_{z_t}^2 = \text{softplus}(W_{\sigma_z} g^{\theta}([z_{t-1}, h_t^e]) + b_{\sigma_z}) + \ln \epsilon_{\sigma_z}^2 \end{cases};$$

Random sampling z from latent space by reparameterization technique: $z_t = \mu_{z_t} + \sigma_{z_t} \cdot \epsilon_{z_t}$;

Update the latent variable of decoder: $h_t^d = (1 - c_t) \otimes h_{t-1}^d + c_t \otimes \tanh(W_h \cdot z_t + U_h \cdot (r_t \cdot h_{t-1}^d) + b_h^d)$;

Decoder calculates the parameters of the probability distribution of the reconstructed data:

$$\begin{cases} \mu_{\hat{x}_t} = W_{\mu_{\hat{x}}} h_t^d + b_{\mu_{\hat{x}}} \\ \sigma_{\hat{x}_t}^2 = \text{softplus}(W_{\sigma_{\hat{x}}} g^{\phi}(h_t^d) + b_{\sigma_{\hat{x}}}) + \epsilon_{\sigma_{\hat{x}}}^2 \end{cases};$$

Decoder generates probability distributions for reconstructed data \hat{x} : $q_{\phi}(\hat{x}|z) = \mathcal{N}(\mu_{\hat{x}_t}, \sigma_{\hat{x}_t}^2 I)$

Optimize encoder and decoder (θ, ϕ) synchronously according to loss function:

$$\mathcal{L}(\theta, \phi; \mathcal{X}_{train}^{(i)}) = \max_{\theta, \phi} \left\{ \mathbb{E}_{z \sim p_{\theta}(z|\mathcal{X}_{train}^{(i)})} [\ln q_{\phi}(\hat{x}|z)] - D_{\text{KL}}[p_{\theta}(z|\mathcal{X}_{train}^{(i)}) || p_{\theta}(z)] \right\};$$

end for

5、异常评分

均方误差（MSE）作为原始向量和重构向量之间的重构误差指标。

$$\tilde{SP}_t = \frac{1}{M} \sum_{m=0}^{M-1} \left(x_t^{(m)} - \hat{x}_t^{(m)} \right)^2$$

其中 t 表示采样时间， m 表示信号维度。对每个窗口中的SP进行平均，得到每个窗口的异常分数（SW）。

$$SW_l = \frac{1}{T} \sum_{t=0}^{T-1} [\tilde{SP}_t] = \frac{1}{T} \sum_{t=0}^{T-1} \left[\frac{1}{M} \sum_{m=0}^{M-1} \left(x_{(l)}^{(t,m)} - \hat{x}_{(l)}^{(t,m)} \right)^2 \right]$$

其中 l 表示 x 中的第 l 个窗口， sw_l 的值被标记到第 l 个窗口的结束时间戳。由于相邻窗口之间存在重叠，因此根据重叠次数对重叠部分的SP进行平均，得到去除滑动窗口后整个时间序列的SP：

$$SP_t = \frac{1}{ol} \sum_1^{ol} \tilde{SP}_t$$

其中 ol 表示采样时间点在相邻窗口中出现的次数。

6、阈值选择

设置一个分数阈值来过滤掉包含潜在异常行为的窗口，利用EVT的POT模型来离线设置异常阈值。POT的基本思想是利用广义帕累托分布（GPD）来拟合概率分布的尾部，即超过阈值的极值分布。它预先设置一个合适的阈值，然后对所有超过该阈值的数据组成的数据组进行建模。对于足够大的阈值 th ，大于 th 部分的条件分布 F 渐近服从GPD：

$$\bar{F}(s) = P(S - th_l > s | S > th_l) \sim \left(1 + \frac{\gamma s}{\beta}\right)^{-\frac{1}{\gamma}}$$

其中 th_l 是初始阈值， γ 和 β 是GPD的形状和尺度参数， s 是目标序列 $\{S_1, S_2, \dots, S_L\}$ 的任意值，其中 th_l 以上部分表示为 $S - th_l$ 。参数 γ 和 β 通过最大似然估计(MLE)方法获得并表示为 $\hat{\gamma}$ 和 $\hat{\beta}$ 。最终阈值 th_F 计算如下：

$$th_F \simeq th_l + \frac{\hat{\beta}}{\hat{\gamma}} \left(\left(\frac{qL}{N_{S-th_l}} \right)^{-\hat{\gamma}} - 1 \right)$$

其中 q 是 $S > th$ 的预期部分， L 是序列的长度， N_{S-th_l} 是超过 th_l 的 S_i 的数量。因此，POT方法中需要调整初始阈值 th_l 和高分位数 q 两个参数来设置阈值。

6、阈值选择

POT算法的原理如下：首先，我们有 n 个观测值 x_1, x_1, \dots, x_n 和初始高经验分位数 q 。然后检索超过 q 的峰值集合 Y_t 并对其拟合GDP。最后，我们可以推断极值的分布并计算阈值 th_F 。

Algorithm 2. POT (Peaks over Threshold)

Input: $\{S_1, S_2, \dots, S_n, q\}$

Output: th_F

Step:

$th_I \leftarrow$ Set initial threshold with $\{S_1, S_2, \dots, S_n\}$

$Y_t \leftarrow \{S_i - th_I | S_i > th_I\}$

$\hat{\gamma}, \hat{\beta} \leftarrow \text{MLE}(Y_t)$

$th_F \leftarrow \text{Calculate Threshold } (q, \hat{\gamma}, \hat{\beta}, n, N_{S-th_I}, th_I)$

四、评估

1、数据

使用了从四辆相同规格的纯电动汽车收集的四个数据集，数据采样间隔为10 s。
根据电池组的运行记录，将这些数据集分为两类：

(1) 将没有记录故障报警的数据集1和数据集2作为训练数据集。

(2) 使用包含一些隐藏警报的数据集3和数据集4作为测试数据集。

Dataset statistics.

Dataset	No. 1	No. 2	No. 3	No. 4
Total points	174 683	167 513	234 479	200 727
Driving points	107 811	147 340	185 062	124 398
Anomaly points	0	0	11	12

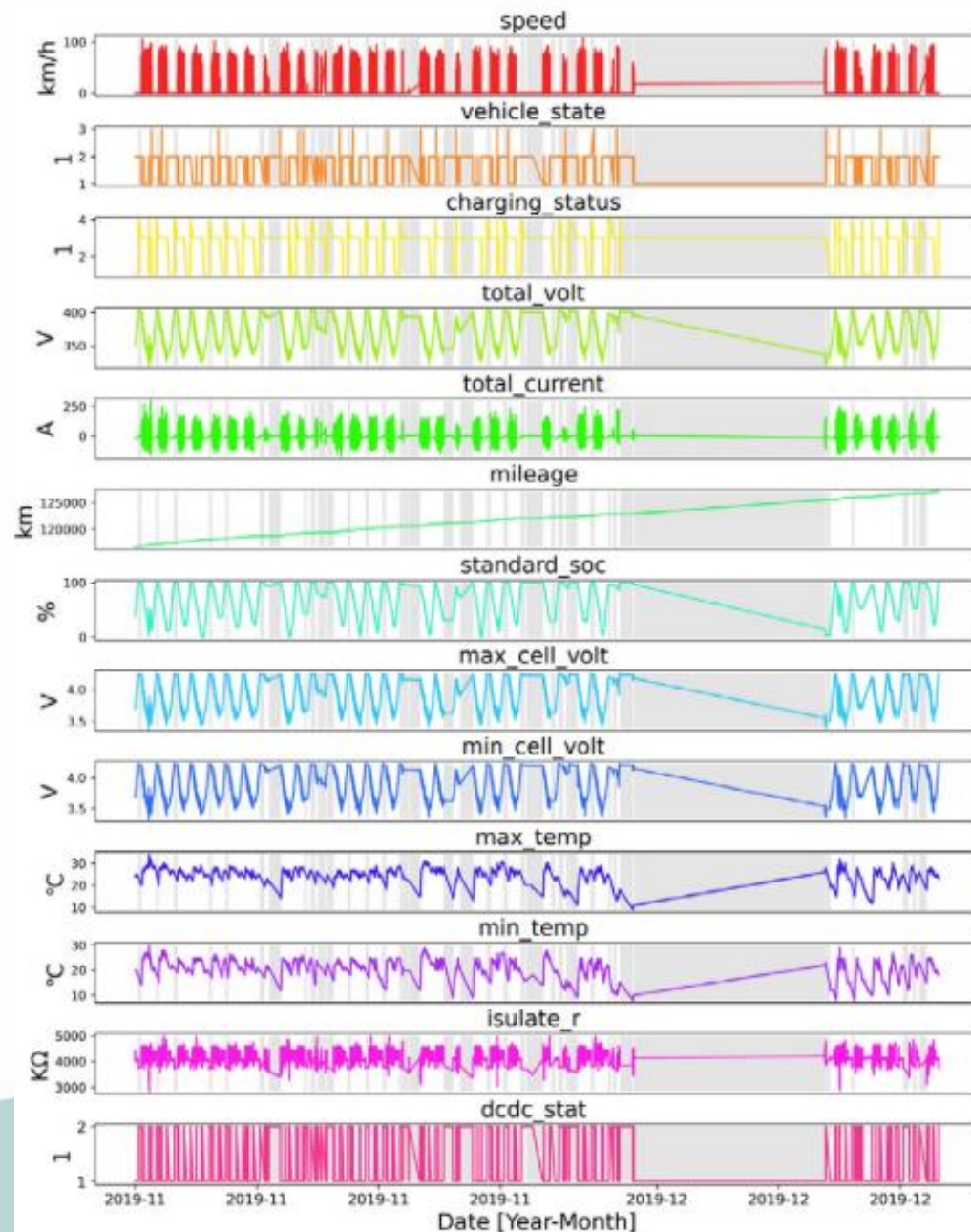


Fig. 5. The 13-dimensional time series of dataset No. 1.

1、数据

从图中可以明显看出，不同维度之间的相关程度是不同的。将所有维度输入GRU-VAE进行训练时，它总是会因为损失不收敛而提前停止。维度之间存在不同程度的相关性，从而导致训练过程中出现过拟合。因此，使用皮尔逊相关系数（PCC）进行相关分析，以选择合适的输入维度（IDs）。

Correlation strength scale													
No correlation		Weak correlation		Moderate correlation		Strong correlation		Extremely strong correlation					
	0	0.2	0.4	0.6	0.8	1							
	speed	vehicle_state	charging_status	total_volt	total_current	mileage	standard_soc	max_cell_volt	min_cell_volt	max_temp	min_temp	isulate_r	dcdc_stat
speed	1.000	0.379	0.345	0.109	0.457	0.094	0.010	0.105	0.113	0.032	0.044	0.160	0.052
vehicle_state	0.379	1.000	0.911	0.168	0.576	0.120	0.012	0.157	0.178	0.099	0.070	0.384	0.137
charging_status	0.345	0.911	1.000	0.174	0.596	0.132	0.019	0.164	0.184	0.114	0.080	0.218	0.088
total_volt	0.109	0.168	0.174	1.000	0.132	0.041	0.959	1.000	1.000	0.058	0.044	0.020	0.025
total_current	0.457	0.576	0.596	0.132	1.000	0.034	0.034	0.123	0.141	0.119	0.099	0.164	0.004
mileage	0.094	0.120	0.132	0.041	0.034	1.000	0.076	0.033	0.049	0.781	0.793	0.004	0.057
standard_soc	0.010	0.012	0.019	0.959	0.034	0.076	1.000	0.961	0.956	0.028	0.021	0.026	0.007
max_cell_volt	0.105	0.157	0.164	1.000	0.123	0.033	0.961	1.000	0.999	0.046	0.033	0.014	0.026
min_cell_volt	0.113	0.178	0.184	1.000	0.141	0.049	0.956	0.999	1.000	0.069	0.055	0.026	0.024
max_temp	0.032	0.099	0.114	0.058	0.119	0.781	0.028	0.046	0.069	1.000	0.995	0.021	0.088
min_temp	0.044	0.070	0.080	0.044	0.099	0.793	0.021	0.033	0.055	0.995	1.000	0.013	0.086
isulate_r	0.160	0.384	0.218	0.020	0.164	0.004	0.026	0.014	0.026	0.021	0.013	1.000	0.381
dcdc_stat	0.052	0.137	0.088	0.025	0.004	0.057	0.007	0.026	0.024	0.088	0.086	0.381	1.000

可以发现四组相关度较高：

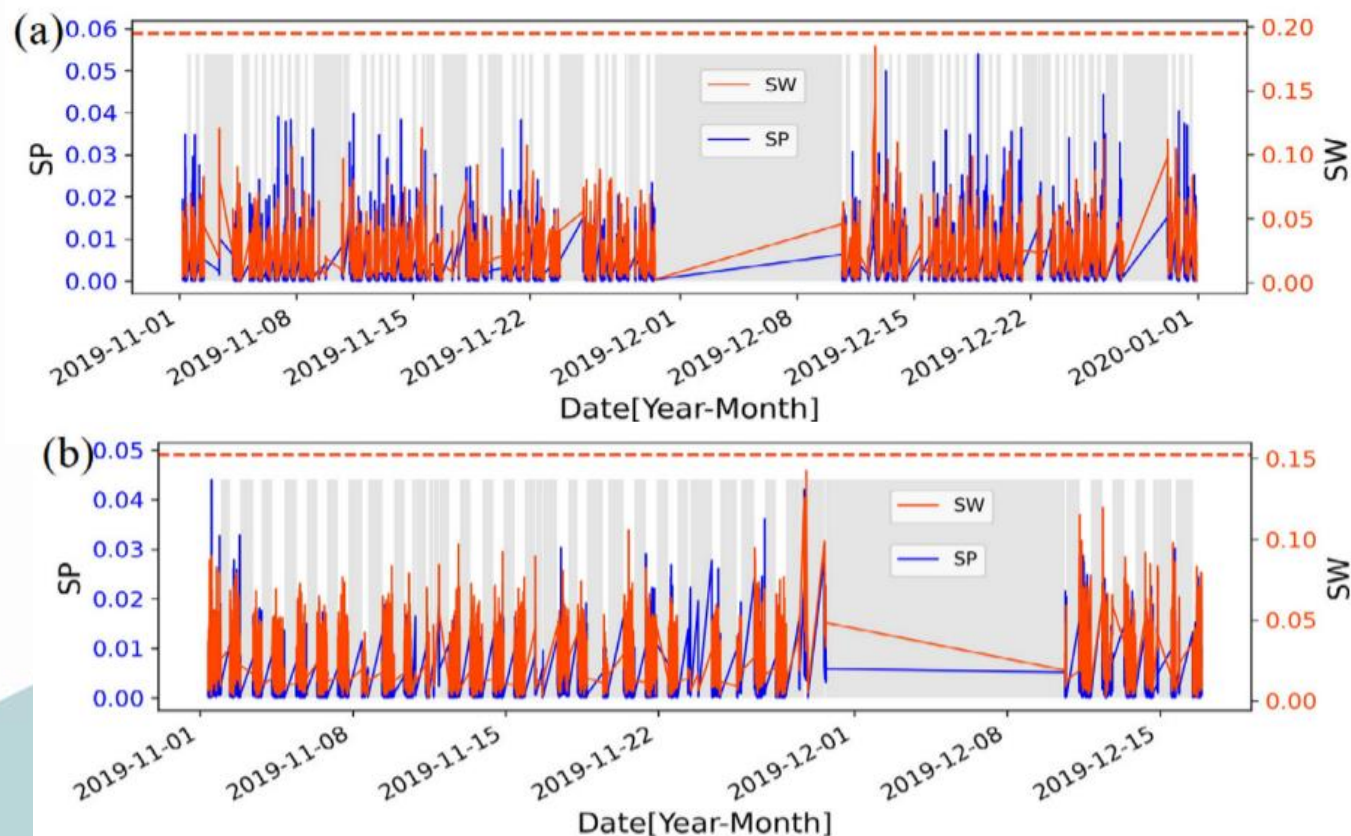
(1) vehicle_state和charging_status；(2)total_volt、standard_soc、max_cell_volt和min_cell_volt；(3) 最大温度和最小温度；(4) isulate_r 和 dcdc_stat。

2、模型训练与验证

由于PCC仅对变量之间的线性关系敏感，因此无法评估瞬时变化较强的速度和total_current。因此，speed和total_current也作为评价模型重构能力的IDs之一。

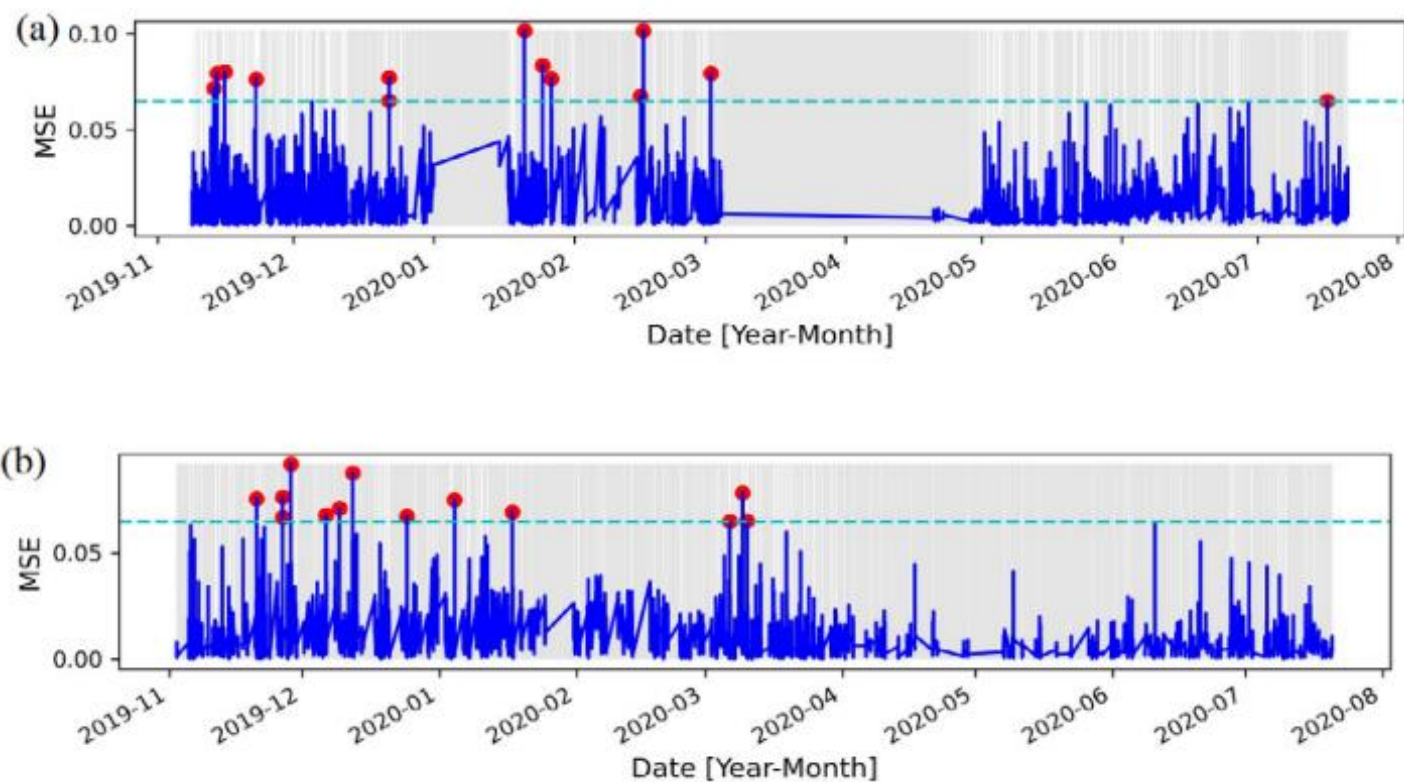
Total_volt、total_current、standard_soc、max_cell_volt、min_cell_volt、max_temp 和 min_temp 作为模型的 IDs。

分别使用1号和2号数据集训练GRU-VAE，得到对应的两组模型参数A和B，然后用于1号和2号数据集的交叉验证。模型参数A对应的SW和SP分别小于0.2和0.06，模型参数B对应的SW和SP分别小于0.15和0.05，为POT的thI选择提供参考。



2、模型训练与验证

使用模型参数 A 来模拟 3 号和 4 号数据集的在线异常检测。基于 thI 预选为 0.06 将高分位数 q 设置为 $1 \sim 10^{-4}$ ，根据计算得到的 SP 最终阈值 thF 为 0.0648。



2、模型训练与验证

为了进一步观察异常的演化过程，提取SP的最大3个值进行部分放大。观察每一个维度再这个时间结点附近的变化情况。

在实践中，异常的持续时间并不固定，要么迅速缓解直至消失，要么趋于恶化并发展为故障。因此，跟踪连续滑动窗口的异常分数的变化有助于了解异常的增长趋势。若异常缓解直至消失，则作为预警提示，若异常持续或加剧，则触发警报。

可以根据序列号统计数据集中记录的产生最大值和最小值的电池块和温度探头的数量。若某几个序列号频繁出现记录，表明电池块的一致性趋于恶化或者探头附近存在热失控的潜在风险。

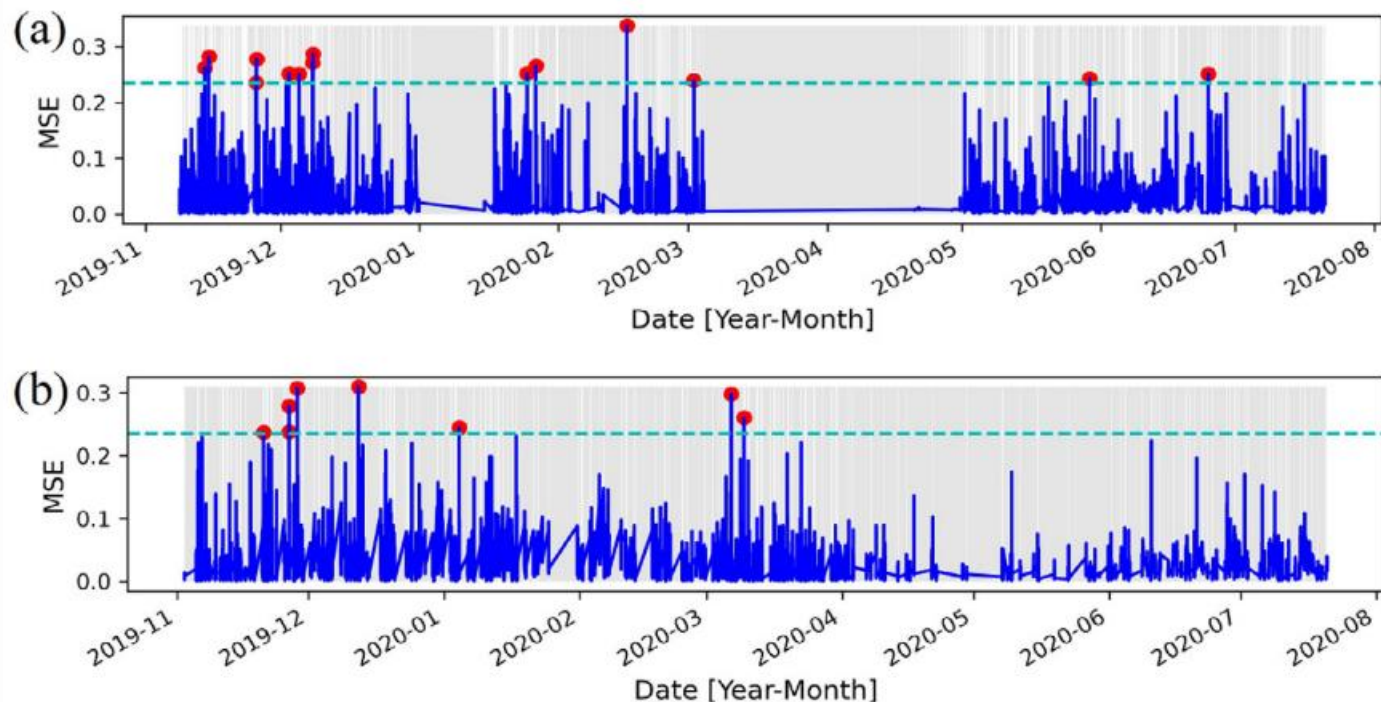
Field name	Number of the brick recorded (Sorted by SP)
<i>max_cell_volt</i>	55, 41, 79, 79, 41, 41, 55, 79, 79, 79, 41, 79, 51
<i>min_cell_volt</i>	13, 13, 13, 13, 25, 13, 13, 13, 13, 13, 13, 13, 13
<i>max_temp</i>	15, 15, 21, 15, 15, 16, 15, 21, 15, 3, 15, 48, 25
<i>min_temp</i>	2, 2, 13, 14, 2, 7, 10, 2, 10, 14, 9, 48, 9

3、模型比较

比较了统一数据集下，GRU-VAE和GRU-AE的异常检测性能。

AE与VAE相比，MSE值整体有所增加，表明AE的重建能力低于VAE。

使用精确率和召回率来评估这两个模型的检测性能。平均而言，GRU-VAE的F1分数为0.818，GRU-AE的F1分数为0.66。



Model	Dataset No. 3		Dataset No. 4	
	Precision	Recall	Precision	Recall
GRU-VAE	11/13	11/11	9/13	9/12
GRU-AE	9/14	9/11	6/8	6/12

五、结论

1、总结

GRU-VAE以集成电池组监测数据和电动汽车运行数据的MVTs为输入，仅使用正常数据进行半监督模型训练，无需先验异常类别，然后通过评估其重构误差来判断测试样本是否包含异常事件。其中，POT模型用于合理设置异常检测阈值。

通过对真实电动汽车运行数据集的训练和测试，检测出潜在不一致的块单元，并且该模型在信号维度和电池组尺寸上具有可扩展性，从而为电动汽车动力电池组的异常检测提供了稳健的模型。

与基于UVTS、针对特定类型故障的异常检测方法相比，该方法为系统级异常检测提供了一种通用的框架，不仅可以定位潜在异常发生的具体时间，还可以识别潜在不一致的电池块，以便提醒BMS及时采取容错措施，防止异常导致进一步劣化。