DOI：10.13234/j.issn.2095-2805.2025.8.293 **中图分类号**：TM912.9 **文献标志码**：A

基于 VAE 模型的 EIS 数据增强的锂电池 SOH 和 RUL 预测

常 伟，胡志超，潘多昭，师继文
(南通乐创新能源有限公司，南通 200233)

摘要：电池的健康状态 SOH(state-of-health)和剩余使用寿命 RUL(remaining useful life)是衡量电池性能和状态的重要指标，对它们的准确预测具有很高的实用价值。电池的电化学阻抗谱 EIS(electrochemical impedance spectroscopy)测试能够揭示电池内部的电化学反应和过程，为评估电池的健康状况提供重要信息。EIS 可用于预测 SOH 和 RUL，但测试成本高、数据量少。采用变分自编码器 VAE(variational autoencoder)进行数据增强，可以扩充 EIS 数据的数量，有助于提升模型的泛化能力并减少过拟合的风险。EIS 数据通常是高维的且可能包含冗余信息，主成分分析 PCA(principal component analysis)可以用来降低数据的维度，同时保留关键信息。卷积神经网络 CNN(convolutional neural network)能够识别 EIS 数据中的空间特征，而双向长短期记忆 BiLSTM(bidirectional long short-term memory)网络能够捕捉时间序列数据中的特征。注意力(Attention)机制则可以加强模型对数据上下文关系的理解。研究结果表明，结合 PCA、CNN、BiLSTM 和 Attention 的模型，在加入增强数据后，能够更有效地预测电池的 SOH 和 RUL。所提 EIS 数据增强方法也可以被应用于其他模型，以提高预测性能。

关键词：变分自编码；阻抗谱；电化学阻抗谱；健康状态；剩余使用寿命；主成分分析；卷积神经网络；双向长短期记忆网络

Prediction of SOH and RUL for Lithium-ion Batteries Using EIS Data Augmentation Based on VAE Model

CHANG Wei, HU Zhichao, PAN Duozhao, SHI Jiwen
(Nantong LeChuang New Energy Co., Ltd., Nantong 200233, China)

Abstract: The state-of-health (SOH) and remaining useful life (RUL) of batteries are important indicators for measuring their performance and conditions, and an accurate prediction of these indicators has high practical value. The electrochemical impedance spectroscopy (EIS) test of batteries can reveal the electrochemical reactions and processes inside them, providing important information for evaluating their SOH. EIS can be used to predict SOH and RUL, but it has high testing costs and limited data volume. The use of variational autoencoder (VAE) for data augmentation can expand the volume of EIS data, which helps improve the model's generalization capability and reduce the risk of overfitting. Due to the fact that the EIS data is usually high-dimensional and may contain redundant information, principal component analysis (PCA) can be used to reduce the dimensionality of data while retaining the key information. Convolutional neural network (CNN) can recognize the spatial features in the EIS data, while bidirectional long short-term memory (BiLSTM) network can capture features within the time series data. The Attention mechanism can enhance the model's understanding of data contextual relationships. The research results indicate that the model combining PCA, CNN, BiLSTM network and Attention mechanism can more effectively predict the SOH and RUL of batteries with the addition of augmented data. The proposed EIS data augmentation method can also be applied to other models to improve the predictive performance.

Keywords: Variational autoencoding; impedance spectrum; electrochemical impedance spectroscopy (EIS); state-of-health (SOH); remaining useful life (RUL); principal component analysis (PCA); convolutional neural network (CNN); bidirectional long short-term memory (BiLSTM) network

收稿日期：2024-01-18；**修回日期**：2024-02-26；**录用日期**：2024-03-27；**网络首发日期**：2024-08-20

电化学阻抗谱 EIS(electrochemical impedance spectroscopy)是 1 种用于分析电池电化学特性的技术，广泛应用于电池健康状态 SOH(state-of-health)和剩余使用寿命 RUL(remaining useful life)的预测。在特征选取上，可以通过 EIS 的实部和虚部数据直接进行 SOH 或 RUL 的预测[1-9]。在特征降维方面，从 EIS 所涵盖的多个频率中，选择具有代表性的频率可以显著降低模型复杂度。然而，代表性频率的选择通常具有主观性，缺乏科学的依据[10]。在模型选择上，传统机器学习模型[11]对于 SOH 的预测具有良好的效果，但其预测精度有待提升。

对电池阻抗谱数据进行测试时，必须依靠专业的实验仪器，并且测试的条件要求严格，导致测试数据的获取成本较高，因此可用的测试数据量通常较少。鉴于已有测试数据数量有限，探索使用数据增强技术来扩充电池阻抗谱数据集具有重要意义[12]。

本研究的创新之处体现在 2 个方面：(1)应用变分自编码器 VAE(variational autoencoder)模型对实际的 EIS 测试数据进行增强，扩大了训练数据集的规模，进而增强了模型在面对未知数据时的适应能力，同时提高了模型对扰动的抵抗力，并有效减少了模型对训练数据过度依赖的风险；(2)在主成分分析-卷积神经网络-双向长短期记忆网络-注意力 PCA-CNN-BiLSTM-Attention(principal component analysis - convolutional neural network - bidirectional long short - term memory network - Attention)模型上，对比了引入 EIS 增强数据前后对电池 SOH 和 RUL 预测效果的差异，结果表明加入增强数据后模型的预测精度显著提高。本文的研究成果为 EIS 数据增强提供了 1 条可行路径，也为 SOH 和 RUL 预测提供了 1 种参考模型。所提出的 EIS 数据增强方法具有普适性，可以应用于其他模型以提升其预测性能。

1 数据采集

在选择和校准 EIS 测试设备时，确保精确度至关重要。例如，可以使用电化学分析测试仪来进行 EIS 测试，这样可以保证所获得的测试数据是准确的。EIS 测试通常会在不同的温度和频率下进行，以便收集到电池在不同条件下的阻抗谱。本文设定测试的温度分别为 25、35、45 ℃，频率范围为 10^{-2}~10^{5} Hz，在这个范围内选取 60 个具体的频率值，频率值的详细信息见表 1。

表 1 EIS 频率取值(Hz)

Tab. 1 Values of EIS frequency (Hz)

第 1 组	第 2 组	第 3 组	第 4 组	第 5 组
20 004.40	1 203.840	72.517 0	4.369 41	0.262 61
15 829.10	952.865	57.368 1	3.456 86	0.207 91
12 516.70	754.275	45.362 9	2.735 47	0.164 52
9 909.44	596.718	35.931 3	2.160 54	0.130 07
7 835.48	471.963	28.409 0	1.709 52	0.103 09
6 217.24	373.208	22.482 0	1.353 52	0.081 53
4 905.29	295.472	17.796 1	1.070 79	0.064 43
3 881.27	233.877	14.068 1	0.847 34	0.051 02
3 070.98	185.059	11.144 8	0.670 72	0.040 42
2 430.77	146.358	8.817 7	0.530 67	0.031 92
1 923.15	115.778	6.975 4	0.419 76	0.025 28
1 522.43	91.672	5.517 3	0.331 83	0.019 99

在进行 EIS 测试时，必须注意保持测试条件(温度、电池状态等)的一致性，这是为了保证测试结果的准确性和可重复性。在数据采集过程中，对电池进行多次完整的充电和放电循环测试，并记录每个频率点对应的阻抗数据，以及相应的电池 SOH 和 RUL 数据。

2 EIS 数据增强

2.1 VAE 模型

VAE 是 1 种生成模型，可以学习数据的隐空间分布，并从中随机采样生成新数据，这种随机生成

过程可以实现数据的增强。VAE 模型结构如图 1 所示，通过在潜在分布进行采样，利用解码器将采样的潜在向量转换成新的数据样本。编码器将输入数据 x 映射到潜在分布 $q(\boldsymbol{z}|x)$中，然后将潜在分布映射到符合均值 μ 和方差 σ 的高斯分布 $N(\mu, \sigma)$中，对高斯分布进行采样，对采样样本加入服从 $N(0.00, 0.01)$的随机扰动，得到采样样本即潜在向量 $\boldsymbol{z}$；解码器将潜在向量 $\boldsymbol{z}$ 还原为 $\hat{x}$，此时解码器中的变量服从分布 $p(\hat{x}|\boldsymbol{z})$。

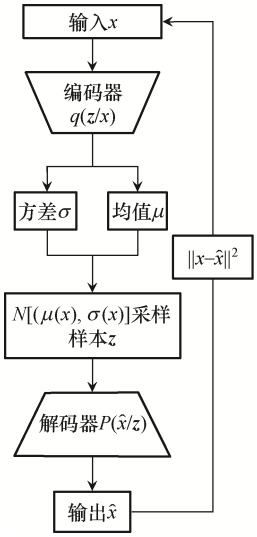


图 1　VAE 模型结构

Fig. 1 Structure of VAE model

2.2 EIS 数据增强模型

基于 VAE 模型的 EIS 数据增强模型结构如图 2 所示。图中，real1~real60 表示阻抗数据的实数部分，imaginary1~imaginary60 表示阻抗数据的虚数部分，两者构成输入数据。首先，输入数据连接到编码器，编码器将输入数据映射到潜在分布 $q(\boldsymbol{z}|x)$中；然后将潜在分布映射到高斯分布中，分别连接至均值 μ 和方差 σ；接着对高斯分布进行采样，加入服从 $N(0.00, 0.01)$的随机扰动，生成采样样本 $\boldsymbol{z}$，即潜在变量；最后将采样得到的潜在变量 $\boldsymbol{z}$ 输入到解码器中，解码器输出预测的数值，即增强数据。本文中，由于预测值为连续数值，VAE 模型中的重构误差为均方误差 MSE (mean square error)，因此整个模型的损失函数为均方误差与潜在变量的 KL 散度(Kullback-Leibler divergence)误差之和。

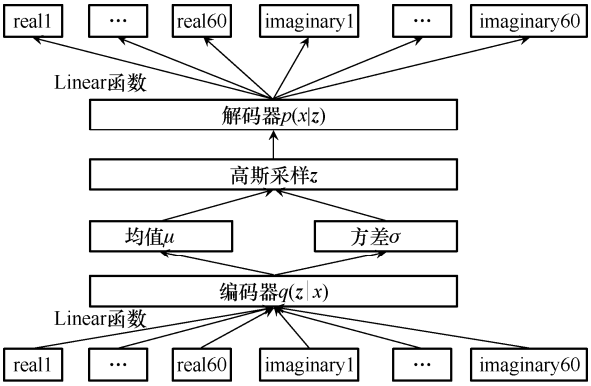


图 2　基于 VAE 模型的 EIS 数据增强模型结构

Fig. 2 Structure of EIS data augmentation model based on VAE model

本文基于 VAE 模型进行的 EIS 数据增强方法的优点在于：(1)VAE 模型可以学习到数据中的潜在变量，因此可以生成全新的样本，而不是简单地重复输入数据；(2)增强数据中，只对输入数据 EIS 进行预测，不对预测目标进行预测，减少了预测误差的发生概率。

3　电池 SOH 和 RUL 预测

3.1 电池 SOH 和 RUL 预测流程

基于 PCA-CNN-BiLSTM-Attention 模型和 EIS 数据预测电池 SOH 和 RUL 的流程图如图 3 所示。第 1 步，通过实验收集电池的 EIS、SOH 和 RUL 的数据；第 2 步，使用 VAE 对 EIS 数据进行增强，以增加数据的多样性和信息量；第 3 步，使用 PCA 对原始 EIS 数据和经过增强的数据进行降维，以提取主要特征；第 4 步，使用 CNN 从 EIS 数据中提取空间特征；第 5 步，使用 BiLSTM 网络提取 EIS 数据中的时间序列特征；第 6 步，使用 Attention 模型识别 EIS 数据中最关键的特征；第 7 步，基于提取出的特征预测电池的 SOH 和 RUL。

3.2 构建 SOH 和 RUL 预测模型

本文使用的 PCA-CNN-BiLSTM-Attention 模型结构如图 4 所示[13]。输入数据由 2 部分组成：real1~real60 表示阻抗数据的实数部分，imaginary1~imaginary60 表示阻抗数据的虚数部分。这些数据



图 3 预测电池 SOH 和 RUL 流程

Fig. 3 Flow chart of prediction of battery SOH and RUL

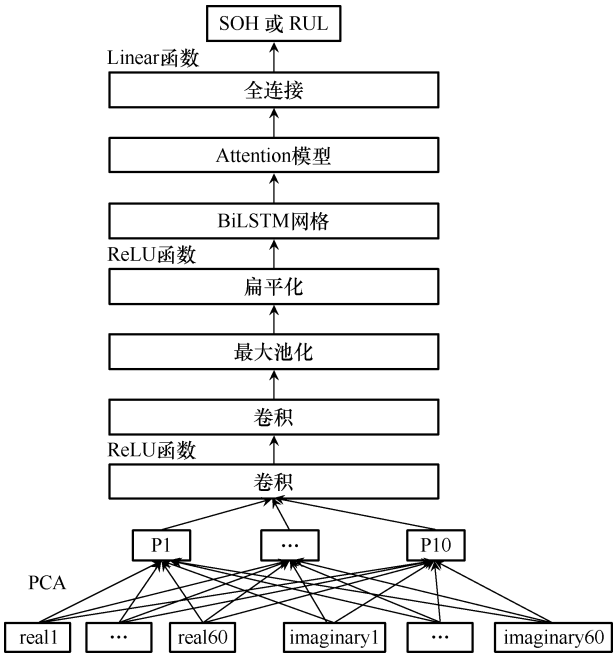


图 4 预测电池 SOH 和 RUL 的模型结构

Fig. 4 Structure of model for predicting battery SOH and RUL

集包含了原始的 EIS 测量结果及经过数据增强处理的数据。P1~P10 为通过对数据进行 PCA 后得到的主要成分，其被输入到 1 个 2 层卷积神经网络，该网络使用 ReLU 作为激活函数，并通过最大池化操作来提取关键特征。然后，将这些特征进行扁平化操作，转换成 1 组一维数组，随后输入到使用 ReLU 作为激活函数的 BiLSTM 结构中。BiLSTM 的输出被送入Attention模型，以帮助模型识别序列中各个时间步的相对重要性。最后，通过 1 个线性回归层来预测电池的 SOH 或 RUL。

3.2.1 PCA 提取电化学阻抗谱主成分

在进行 EIS 测试时，每个测试频率都会产生 1 个包含实部和虚部的阻抗值，总共形成 120 个数据点作为分析的基础。但是，这 120 个数据点并非都对评估电池的 SOH 有贡献；某些频率下的数据可能与预测目标无关，或者含有噪声，可能会对预测结果的准确性产生负面影响。因此，为了提升预测的精确度，输入数据可以通过降维来保留对预测电池 SOH 最有用的特征。

PCA 是 1 种线性降维方法[13]，通过正交变换将数据从高维降至低维，同时保留大部分信息。图 5 展示了 PCA 的降维过程，其中圆圈代表样本点，在 *XY* 平面上分布，其分散性由变量方差衡量。PCA 通过线性转换将数据映射至新坐标系，使数据在此坐标系中方差最大。

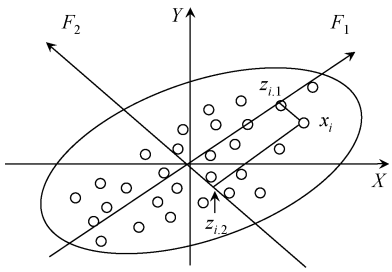


图 5 PCA 几何解释示意

Fig. 5 Schematic of geometric interpretation for PCA

PCA 将原始数据转换为 1 组新的相互正交的特征向量，这些特征向量被称为主成分。PCA 的目标是在减少数据维度的过程中，尽可能地保持数据的信息完整性。

在本研究中，通过 PCA 提取了 EIS 的 10 个主成分，它们共同解释了输入数据中 99%的方差，实现了高效率的数据降维，同时几乎 100%保留了原始信息。

3.2.2 CNN 模型提取 EIS 数据的时空特征

CNN 能够自动识别输入数据的局部特征并提

取更高层次的抽象信息。本研究设计了 1 个 CNN 模型，包含 2 个卷积层和 1 个池化层[13]，在每个卷积层使用了 ReLU 激活函数，并配备了 32 个过滤器，每个过滤器的核大小为 1。此外，池化层同样采用了大小为 1 的池化窗口。这种结构设计有助于网络从输入数据中学习到有效的特征标识[13]。

图 6 为卷积计算示意图，输入的 6×6 特征图经过 3×3 卷积核处理，生成 4×4 的输出特征图[13]。这个过程通过计算卷积核与输入特征图对应区域的张量积并移动卷积核完成。池化层通过最大池化或平均池化减少特征图尺寸，如图 7 所示，本文采用最大池化。最终，通过扁平化层将多维特征转换为一维向量，作为 EIS 数据的空间特征表示。

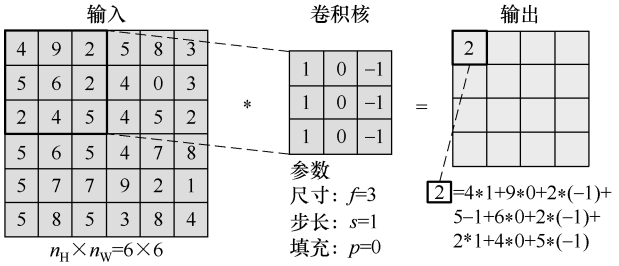


图 6　卷积计算示意

Fig. 6 Schematic of convolution calculation

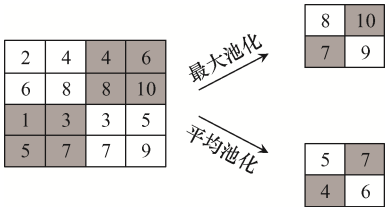


图 7　池化计算示意

Fig. 7 Schematic of pooling calculation

在本文中，首先使用 PCA 模型对原始 EIS 数据和经过增强的 EIS 数据进行降维处理，然后将这 2 组降维后的数据分别以 8:2 的比例划分为训练集和测试集。接着，将这 2 套训练集合并，作为 CNN 模型的输入数据。针对回归预测模型，对每 1 行数据，即每个一维向量，进行卷积和池化计算[13]。

3.2.3　BiLSTM 模型提取 EIS 的时间序列变化规律

BiLSTM 模型由前向和后向 2 个 LSTM 网络组成。LSTM 模型结构如图 8 所示，其主要构成包括 6 部分[13]：①输入门，负责决定当前输入的哪些信息可以传递到状态单元中；②遗忘门，负责决定状态单元中的哪些信息应该被丢弃；③状态单元，是保存历史信息的关键部分；④输出门，负责决定状态单元中的哪些信息应该被传递到下一个隐藏层；⑤激活函数，用于进行非线性处理；⑥输出层，与输入数据相对应，负责生成 LSTM 的最终输出。

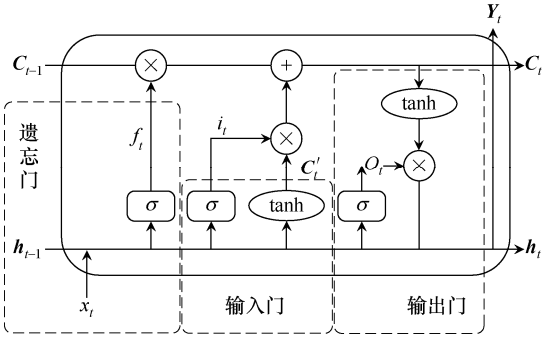


图 8　LSTM 模型结构

Fig. 8 Structure of LSTM model

BiLSTM 模型由前向和后向 LSTM 单元构成，每个单元处理序列的不同方向。这 2 个方向的隐藏状态随后合并，产生模型的输出状态[13]，如图 9 所示。输入数据 X，通过拼接前向 LSTM 的隐藏状态 h_i^{f} 和后向 LSTM 的隐藏状态 h_{n-i}^{b}，实现在第 i 个时间步输出 $\boldsymbol{h}_i=[h_i^{\mathrm{f}}\cdot h_{n-i}^{\mathrm{b}}]$[13]。

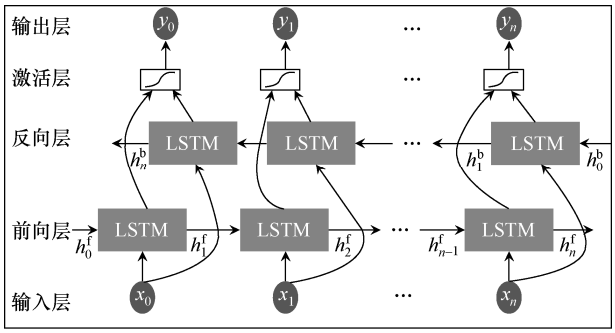


图 9　BiLSTM 模型结构

Fig. 9 Structure of BiLSTM model

CNN-BiLSTM 模型结合了 CNN 和 BiLSTM，能够从 EIS 降维后的数据中提取空间和时间维度的特征，即时空特征。

3.2.4　Attention 模型提取 EIS 时空特征的重要部分

CNN-BiLSTM 模型成功提取了 EIS 数据的时空特征。这些特征随后被送入 Attention 模型，

Attention 模型的作用是评估这些特征的重要性，并为重要的特征分配更高的权重，这样就能强调出特征中最为关键的部分[13]。

本文采用的 Attention 机制基于 Bahdanau Attention 模型，其运作流程包括训练权重向量、应用权重调整序列隐藏状态的贡献度、聚焦关键信息的表征。Bahdanau Attention 模型的具体结构如图 10 所示。

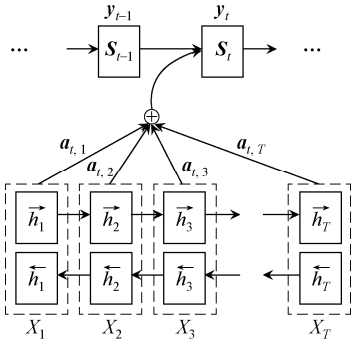


图 10 Bahdanau Attention 模型结构

Fig. 10 Structure of Bahdanau Attention model

Bahdanau Attention 模型分为 Encoder 和 Decoder 2 个计算模块。

(1)Encoder 模块可表示为

$$\boldsymbol{h}_i = \tanh(\boldsymbol{W}[\boldsymbol{h}_{i-1}, x_i]) \tag{1}$$

$$\boldsymbol{o}_i = \mathrm{softmax}(\boldsymbol{V}\boldsymbol{h}_i) \tag{2}$$

式中：$\boldsymbol{h}_i$ 为 i 时刻输入 x_i 的隐藏状态向量；$\boldsymbol{o}_i$ 为 i 时刻隐藏状态的归一化输出向量；$\boldsymbol{W}$ 和 $\boldsymbol{V}$ 为待学习的权重矩阵。

CNN-BiLSTM 模型已经从 EIS 数据中提取出了对应的时空特征。接下来，将 CNN-BiLSTM 的输出结果用作 Attention 模型的输入，即 Encoder 模块中的 x_i。

(2)Decoder 模块分 2 步完成。

第 1 步，生成 t 时刻的语义向量，即

$$\boldsymbol{c}_t = \sum_{i=1}^{T} \boldsymbol{\alpha}_{ti} \boldsymbol{h}_i \tag{3}$$

$$\boldsymbol{\alpha}_{ti} = \frac{\exp(\boldsymbol{e}_{ti})}{\sum_{k=1}^{T} \exp(\boldsymbol{e}_{tk})} \tag{4}$$

$$\boldsymbol{e}_{ti} = \boldsymbol{V}_{\mathrm{a}}^{\mathrm{T}} \tanh(\boldsymbol{W}_{\mathrm{a}}[\boldsymbol{s}_{t-1}, \boldsymbol{h}_i]) \tag{5}$$

式中：$\boldsymbol{c}_t$ 为 t 时刻的语义向量；$\boldsymbol{e}_{ti}$ 为编码器中第 i 时刻隐层状态向量 $\boldsymbol{h}_i$ 对解码器中第 t 时刻隐层状态 $\boldsymbol{s}_t$ 的影响程度；$\boldsymbol{\alpha}_{ti}$ 为通过 softmax 函数将 $\boldsymbol{e}_{ti}$ 归一化之后的值；$\boldsymbol{W}_{\mathrm{a}}$ 为待学习的权重矩阵；$\boldsymbol{V}_{\mathrm{a}}$ 为待学习的权重向量。

第 2 步，传递隐层信息并预测，即

$$\boldsymbol{s}_t = \tanh(\boldsymbol{W}[\boldsymbol{s}_{t-1}, \boldsymbol{y}_{t-1}, \boldsymbol{c}_t]) \tag{6}$$

$$\boldsymbol{y}_t = \mathrm{softmax}(\boldsymbol{V}\boldsymbol{s}_t) \tag{7}$$

式中，$\boldsymbol{y}_t$ 为注意力模型根据特征的相对重要性为其分配的权重，最终输出加权的特征向量[13]。

3.2.5 预测电池 SOH 或 RUL

在注意力模型处理后，模型添加了 1 个全连接层，这个层会进一步从 Attention 模型的输出中提取特征，以便得到时间序列的高级特征表示。这个全连接层有 16 个神经元，每个神经元都使用 ReLU 作为激活函数。

在全连接层后是模型的最后一层，即输出层，它是 1 个线性回归层，专门用于预测电池的 SOH 或 RUL。输出层使用 Linear 作为激活函数，这样可以保证模型的输出是 1 个连续的数值。

在训练过程中，模型采用了 MSE 作为损失函数，该函数通过计算预测值与实际目标值之间差异平方的平均值来衡量模型的准确性。为了优化模型参数，选择 RMSProp 优化算法，该算法适合于递归神经网络。在评估模型的性能时，使用平均绝对误差 MAE(mean absolute error)指标，通过求取目标值与预测值之间差的绝对值之和来评价模型的预测能力。

4 预测效果展示

本文通过公开的数据集合[14]进行了模型效果检验，利用该数据集中的阻抗数据分别对电池的 SOH 和 RUL 进行预测。测试数据包含了 120 个特征，这些特征的编号为 f1~f120。其中，f1~f60 表示阻抗数据的实部，f61~f120 表示阻抗数据的虚部，二者均按频率从高到低排序。

4.1 预测数据展示

4.1.1 预测 SOH 的 EIS 原始数据和增强数据展示

基于 VAE 模型对真实的 EIS 数据进行增强，数据增强模型的整体均方根误差为 0.184 7，误差较小，数据还原效果较好。预测 SOH 的 EIS 数据集合样例(部分)见表 2，其中，f1~f120 表示原始 EIS 阻抗数据，f1_aug~f120_aug 表示增强后的 EIS 阻抗数据。

表 2　预测 SOH 的 EIS 数据集合样例(部分)

Tab. 2 Sample of EIS data set for predicting SOH (part)

f1/Ω	f1_aug/Ω	f2/Ω	f2_aug/Ω	f120/Ω	f120_aug/Ω	SOH/(A·h)
0.384 70	0.374 499 77	0.391 56	0.380 515 87	0.327 95	0.335 009 43	37.20
0.388 86	0.381 822 82	0.393 43	0.388 099 46	0.323 00	0.334 569 40	36.22

4.1.2 预测 RUL 的 EIS 原始数据和增强数据展示

基于 VAE 模型对真实的 EIS 数据进行增强，数据增强模型的整体均方根误差为 0.180 0，误差较小，数据还原效果较好。值得注意的是，用于预测 SOH 和 RUL 的阻抗谱 EIS 数据并不完全相同。预测 RUL 的 EIS 数据集合样例(部分)见表 3。

表 3　预测 RUL 的 EIS 数据集合样例(部分)

Tab. 3 Sample of EIS data set for predicting RUL (part)

f1/Ω	f1_aug/Ω	f2/Ω	f2_aug/Ω	f120/Ω	f120_aug/Ω	RUL/次
0.384 70	0.379 538 77	0.391 56	0.384 051 44	0.327 95	0.332 721 83	234
0.752 11	0.769 263 80	0.760 74	0.776 251 90	0.291 80	0.293 092 94	0

4.2 PCA 提取 EIS 主成分

4.2.1 PCA 提取预测 SOH 的 EIS 主成分

分别对 EIS 原始数据和增强数据的 120 维度数据进行 PCA 计算，得到 10 个主要成分，样例见表 4。其中，P1~P10 表示 EIS 原始数据的 10 个主成分，P1_aug~P10_aug 表示 EIS 增强数据的 10 个主成分。

表 4　预测 SOH 的电池 EIS 原始数据和增强数据 PCA 计算结果样例(部分)

Tab. 4 Sample of original battery EIS data and PCA calculation results of augmented data for predicting SOH (part)

P1	P1_aug	P2	P2_aug	P10	P10_aug
−0.825 7	0.509 4	0.497 9	−0.015 6	0.007 3	0.077 1
−0.766 7	0.543 1	0.480 0	−0.074 1	0.005 1	0.057 4

4.2.2 PCA 提取预测 RUL 的 EIS 主成分

对预测电池 RUL 的 EIS 原始数据和增强数据的 120 维度分别进行 PCA，提取出 10 个主成分，示例见表 5。

表 5　预测 RUL 的电池 EIS 数据 PCA 计算结果样例(部分)

Tab. 5 Sample of PCA calculation results of battery EIS data for predicting RUL (part)

P1	P1_aug	P2	P2_aug	P10	P10_aug
−0.961 4	−0.731 9	0.063 5	−0.604 5	0.009 0	0.002 5
−0.906 7	−0.732 2	0.092 5	−0.650 4	0.003 4	0.000 3

将 EIS 的真实数据和增强数据的 PCA 模型计算结果，均按照 8:2 的比例分成训练集合和测试集合。在生成训练集合和测试集合时，以上 2 套数据按照相同的数据分割方式进行分割，即随机数生成器的随机种子保持相同，保证训练数据不会发生泄漏。在真实数据和增强数据的训练集合的并集上训练预测模型，在真实数据的测试集合上测试效果。

4.3 预测结果展示

4.3.1 不同数据模型效果对比

在 SOH 预测中，只通过真实 EIS 数据进行训练，预测集合和测试集合的均方根误差为 0.146 8 A·h，而加入增强数据之后，模型的均方根误差减少为 0.123 8 A·h；在 RUL 预测中，只通过真实 EIS 数

据进行训练，预测集合和测试集合的均方根误差为 3.303 0 次，而加入增强数据之后，模型的均方根误差减少为 2.901 9 次，说明增强数据可以提升模型的预测精度。对比结果见表 6。

表 6 不同数据模型效果对比

Tab. 6 Comparison of effect among models using different data

所用数据	SOH 预测模型均方根误差/(A·h)	RUL 预测模型均方根误差/次
EIS 数据	0.146 8	3.303 0
EIS 数据+EIS 增强数据	0.123 8	2.901 9

4.3.2 预测样例展示

在 SOH 的测试数据集中，平均值为 30.577 3 A·h，标准差为 4.066 1 A·h；在预测数据集中，平均值为 30.599 8 A·h，标准差为 4.013 2 A·h。对于 RUL 的测试数据集，平均值为 147.828 6 次，标准差为 103.062 1 次；在预测数据集中，平均值为 147.822 4 次，标准差为 102.230 4 次。这些统计数据表明，预测数据与实际数据的均值和标准差都比较接近，2 组数据的分布状态相近，说明模型的预测效果良好。

预测样例见表 7，效果图如图 11 和图 12 所示。从表 7、图 11 和图 12 可以看出，SOH 和 RUL 的预测误差均较小，预测效果良好。

表 7 预测样例(部分)

Tab. 7 Sample of prediction results (part)

参数	样例 1	样例 2	样例 3	样例 4
SOH 真实值/(A·h)	29.16	24.75	27.31	28.09
SOH 预测值/(A·h)	29.35	25.02	27.35	28.21
SOH 预测误差/(A·h)	−0.19	−0.27	−0.04	−0.12
RUL 真实值/次	30	204	136	258
RUL 预测值/次	32	206	134	261
RUL 预测误差/次	2	2	−2	3

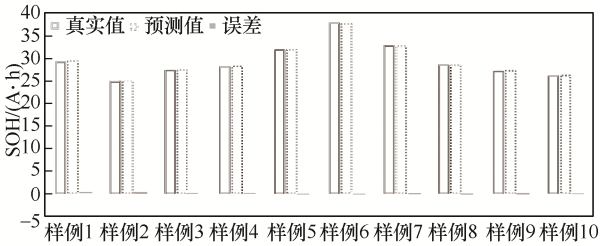


图 11 电池 SOH 预测结果样例

Fig. 11 Sample of battery SOH prediction results

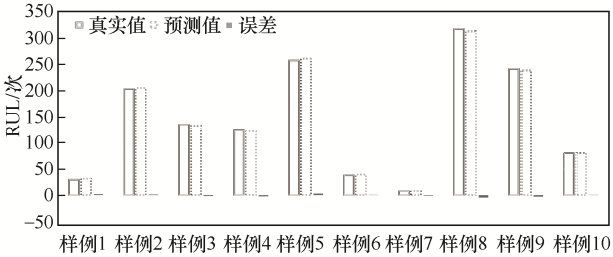


图 12 电池 RUL 预测结果样例

Fig. 12 Sample of battery RUL prediction results

5 结语

本文使用 VAE 模型对 EIS 数据进行增强，使用 PCA 模型提取 EIS 数据的主要特征成分，然后在 CNN-BiLSTM-Attention 模型上对比了使用原始 EIS 和加入 EIS 增强数据之后的 SOH 和 RUL 预测效果。结果表明，加入 EIS 增强数据之后，模型效果有明显提升。本文提出的 EIS 数据增强方法可以推广至其他的预测模型中使用。

参考文献：

[1] 成燕, 郑林弘, 刘江. 基于模态分解和时间序列的锂电池健康状态估计[J]. 电源学报, 2025, 23(5): 241-250.
Cheng Yan, Zheng Linhong, Liu Jiang. Estimation of state-of-health of lithium-ion batteries based on modal decomposition and time series [J]. Journal of Power Supply, 2025, 23(5): 241-250. (in Chinese)

[2] 张帅涛, 蒋品群, 宋树祥, 等. 基于注意力机制和 CNN-LSTM 融合模型的锂电池 SOC 预测[J]. 电源学报, 2024, 22(5): 269-277.
Zhang Shuaitao, Jiang Pinqun, Song Shuxiang, et al. SOC prediction for lithium battery based on fusion model of Attention mechanism and CNN-LSTM [J]. Journal of

Power Supply, 2024,22(5):269-277. (in Chinese)

[3] 贾先屹, 王顺利, 曹文, 等. 基于二阶 RC 等效电路和 SR-DUKF 算法的锂电池 SOC 估算研究[J]. 电源学报, 2024, 22(4): 236-242.
Jia Xianyi, Wang Shunli, Cao Wen, et al. Research on estimation of lithium battery SOC based on second-order RC equivalent circuit and SR-DUKF algorithm [J]. Journal of Power Supply, 2024, 22(4): 236-242. (in Chinese)

[4] 贺伟, 马鸿雁, 张英达, 等. 基于改进门控循环单元神经网络的锂电池组荷电状态预测[J]. 科学技术与工程, 2023, 23(12): 5102-5109.
He Wei, Ma Hongyan, Zhang Yingda, et al. State of charge prediction of the lithium-ion battery pack based on improved gated recurrent unit [J]. Science Technology and Engineering, 2023, 23(12): 5102-5109. (in Chinese)

[5] 徐德鸿, 陈文杰, 何国锋, 等. 新能源对电力电子提出的新课题[J]. 电源学报, 2014, 12(6): 4-9, 30.
Xu Dehong, Chen Wenjie, He Guofeng, et al. New power electronics topics brought by the development of renewable energy [J]. Journal of Power Supply, 2014, 12(6): 4-9, 30. (in Chinese)

[6] 郑永飞, 文怀兴, 韩昉, 等. 基于电池外特征的粒子群神经网络电池健康状态预测[J]. 科学技术与工程, 2019, 19(36): 184-189.
Zheng Yongfei, Wen Huaixing, Han Fang, et al. Prediction of state of health based on particle swarm neural network with battery external characteristics [J]. Science Technology and Engineering, 2019, 19(36): 184-189. (in Chinese)

[7] 白俊琦, 贤燕华. 基于 BP−多时间尺度 UKF 优化算法的 SOC 估计[J]. 电源学报, 2025, 23(3): 363-372.
Bai Junqi, Xian Yanhua. State-of-charge estimation based on BP-multi-timescale UKF optimization algorithm [J]. Journal of Power Supply, 2025, 23(3): 363-372. (in Chinese)

[8] 周雅夫, 万伟东, 董启超. 基于改进多频正弦激励的燃料电池电化学阻抗谱测量[J]. 科学技术与工程, 2022, 22(22): 9679-9685.
Zhou Yafu, Wan Weidong, Dong Qichao. Electrochemistry impedance spectrum measurement of fuel cell based on improved multisine excitation [J]. Science Technology and Engineering, 2022, 22(22): 9679-9685. (in Chinese)

[9] Zhang Yunwei, Tang Qiaochu, Zhang Yao, et al. Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning [J]. Nature Communications, 2020, 11: 1706.

[10] 耿萌萌, 范茂松, 杨凯, 等. 基于 EIS 和神经网络的退役电池 SOH 快速估计[J]. 储能科学与技术, 2022, 11(2): 673-678.
Geng Mengmeng, Fan Maosong, Yang Kai, et al. Fast estimation method for state-of-health of retired batteries based on electrochemical impedance spectroscopy and neural network [J]. Energy Storage Science and Technology, 2022, 11(2): 673-678. (in Chinese)

[11] 刘伟霞, 田勋, 肖家勇, 等. 基于混合模型及 LSTM 的锂电池 SOH 与剩余寿命预测[J]. 储能科学与技术, 2021, 10(2): 689-694.
Liu Weixia, Tian Xun, Xiao Jiayong, et al. Estimation of SOH and remaining life of lithium batteries based on a combination model and long short-term memory [J]. Energy Storage Science and Technology, 2021, 10(2): 689-694. (in Chinese)

[12] Vivier V, Orazem M E. Impedance analysis of electrochemical systems [J]. Chemical Reviews, 2022, 122(12): 11131-11168.

[13] 南通乐创新能源有限公司. 一种基于多模型组合的使用电池 EIS 预测 SOH 的方法: CN117054888A[P]. 2023-11-14.

[14] Zhang Y. ML-identify-battery-degradation [EB/OL]. (2020-01-20)[2024-01-10]. https://github.com/YunweiZhang/ML-identify-battery-degradation.



常伟

作者简介：

常伟(1976—)，男，通信作者，硕士，高级工程师。研究方向：新能源电池算法。E-mail：weichang008@hotmail.com。

胡志超(1984—)，男，硕士，工程师。研究方向：电池热失控，EIS、SOH 估计。E-mail：18221123045@163.com。

潘多昭(1991—)，男，硕士，工程师。研究方向：储能与能源工程，AI 算法，智慧能源。E-mail：pandz_sjtu@163.com。

师继文(1990—)，男，硕士，工程师。研究方向：储能与能源工程，新能源材料。E-mail：sjw_1007@163.com。

(责任编辑：郭凯凯)