

白红艳,何璐瑶,武梓涵,等. 基于 PCA-Adam-LSTM 的涡轮系统状态预测模型[J]. 智能计算机与应用,2026,16(3):142-146. DOI:10.20169/j.issn.2095-2163.24052001

## 基于 PCA-Adam-LSTM 的涡轮系统状态预测模型

白红艳,何璐瑶,武梓涵,杨理践,刘斌

(沈阳工业大学信息科学与工程学院,沈阳 110870)

**摘要:** 涡轮是发动机的重要组成部分,对发动机的整体性能和效率都有重要影响。但传统数据驱动状态预测模型存在输入参数规模过大导致运算负担过重,潜在“鞍点”的问题,难以应对复杂系统的工作场景。因此,为实现对涡轮状态准确预测,保证涡轮健康运行,提出一种基于 PCA-Adam-LSTM 的数据驱动涡轮状态预测方法。建立以 PCA 降维后的 3 维数据为输入,以  $T50$ 、 $P50$  和  $W50$  为输出的数据驱动涡轮状态预测模型。最后,与基于 PCA-BP 和 PCA-LSF 的模型进行对比分析可知,基于 PCA-Adam-LSTM 模型具有准确性更高、泛化能力强和能更好地处理长期依赖关系的优点,平均绝对百分比误差均小于 0.182,平均绝对误差均小于 1.010,均方根误差均小于 1.287。结果表明,基于 PCA-Adam-LSTM 算法对涡轮状态预测有较强实用性和较高准确率,为后续开展涡轮健康度评价提供了理论依据和技术支持。

**关键词:** 涡轮;主成分分析;Adam;数据驱动;LSTM

中图分类号: TG115.28

文献标志码: A

文章编号: 2095-2163(2026)03-0142-05

### PCA-Adam-LSTM-based turbine state prediction model

BAI Hongyan, HE Luyao, WU Zihan, YANG Lijian, LIU Bin

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

**Abstract:** Turbines are vital components of engines, exerting significant influence on overall engine performance and efficiency. However, traditional data-driven state prediction models face challenges such as excessive computational burden due to large input parameter scales and the issue of potential "saddle points," making them ill-suited for dealing with complex systems. To achieve accurate prediction of turbine states and ensure their healthy operation, a data-driven turbine state prediction method based on PCA-Adam-LSTM is proposed. This method establishes a data-driven turbine state prediction model with 3-dimensional input data reduced by PCA and  $T50$ ,  $P50$ , and  $W50$  as output parameters. Finally, a comparative analysis is conducted with models based on PCA-BP and PCA-LSF, revealing that the PCA-Adam-LSTM model exhibits higher accuracy, stronger generalization ability, and better handling of long-term dependencies. The average absolute percentage error is less than 0.182, the average absolute error is less than 1.010, and the root mean square error is less than 1.287. These results demonstrate that the PCA-Adam-LSTM algorithm offers practicality and high accuracy in turbine state prediction, providing a theoretical basis and technical support for subsequent turbine health assessments.

**Key words:** turbine; Principal Component Analysis; Adam; data-driven; LSTM

## 0 引言

发动机健康监测技术通过在发动机上安装传感器收集数据,并利用先进算法实时监控、预测和维护发动机运行状态<sup>[1]</sup>,从而提高飞行器的安全性和可靠性,减少停机时间,提高运行效率,并降低维修成本<sup>[2]</sup>。涡轮是发动机的重要部件,其故障会造成严重后果,因此监测涡轮运行状态非常重要。

传统的部件建模方法基于数学模型<sup>[3-5]</sup>,但通常情况下的输出偏差却会受到多种不确定因素、如生产误差和材料特性<sup>[6]</sup>的影响。在建模时对这些因素将难以进行全面考虑和整合,导致模型与实际情况存在差异。因此,建立数据驱动的涡轮状态预测模型具有重要意义。

近年来,神经网络在发动机部件建模中应用广泛,但未充分考虑输入参数规模大、变量间耦合性

作者简介: 何璐瑶(1994—),女,博士,副教授,主要研究方向:数据驱动建模,人工智能。

通信作者: 白红艳(1996—),女,硕士研究生,主要研究方向:数据驱动建模,人工智能。Email:758269768@qq.com。

收稿日期: 2024-05-20

强、非线性等特点,导致模型运算时间长且负担重。此外,传统神经网络的随机梯度下降法容易陷入“鞍点”,限制了模型优化能力。

本文提出了一种基于 PCA-Adam-LSTM 的方法来建立发动机涡轮关键参数预测模型。首先,应用主成分分析(PCA)技术对大量参数数据进行降维,确保预测准确性的同时减少模型训练和优化时间<sup>[7]</sup>。然后,将降维后的数据集划分为训练集和测试集,并进行归一化操作。最后,采用自适应矩估计优化方法(Adam)作为 LSTM 网络的学习率更新方法,避免了随机梯度下降法的局部最优解问题,并省去手动调整学习率的繁琐过程,实现了对涡轮出口温度、压力和流量的预测。

### 1 PCA-Adam-LSTM 涡轮预测模型

在本研究中,提出一种基于 PCA-Adam-LSTM 方法的数据驱动涡轮状态预测模型,模型的总体框架如图 1 所示,该模型主要包括 3 部分:

- (1)PCA 降维模块。通过 PCA 降维可以在确保精度的同时,缩短模型的运行时间。
- (2)数据预处理模块。将 PCA 降维后数据划分成训练集和测试集,并进行归一化。
- (3)Adam-LSTM 模型训练模块。使用 Adam 优化后的 LSTM 来捕捉涡轮特征信息,输出涡轮状态数据。

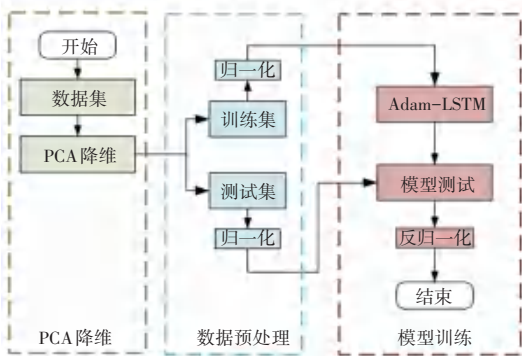


图 1 PCA-Adam-LSTM 涡轮模型流程图

Fig. 1 Flowchart of the PCA-Adam-LSTM turbine model

其中,PCA 是一种用于多变量数据集的统计技术,可通过识别数据中的模式来降低数据的维度。通过 PCA,可以从原始数据中提取关键特征,并构建一组不相关的新的变量<sup>[8]</sup>,这些变量被称为主成分。这些主成分能够捕捉到原始数据集中的大部分信息,从而简化数据结构,实现降维的目的<sup>[9-11]</sup>。

图 2 展示了 LSTM 的单元架构<sup>[12-13]</sup>。LSTM<sup>[14-16]</sup>是循环神经网络(RNN)的一种创新改

进<sup>[17-19]</sup>。与标准 RNN 相比,LSTM 引入了记忆单元,包括遗忘机制、输入机制和输出机制。这些机制显著缓解了 RNN 在处理长序列数据时遇到的梯度爆炸和梯度消失问题<sup>[20-21]</sup>。

图 2 中,  $h$  表示隐藏状态,  $X$  表示输入数据,  $C$  表示单元状态,  $t$  表示时刻,  $\sigma$  表示 Sigmoid 函数,  $\tanh$  表示 tanh 激活函数。

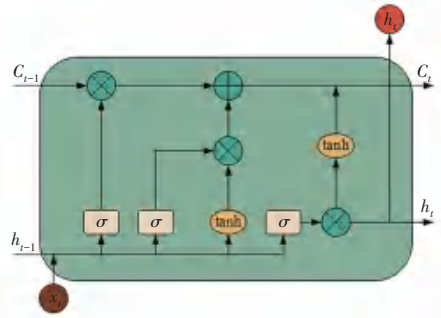


图 2 LSTM 网络单元结构

Fig. 2 Structure of the LSTM network unit

### 2 算例分析

数据集来源于 NASA、新的 CMAPSS 公开数据集(N-CMPASS)。选取 N-CMPASS 中的 DS01\_unit1.0\_cycle2.0 部分,共 4 660 组数据。数据集包含了商用飞机的完整飞行过程数据,涵盖了爬升、巡航和下降飞行过程。

#### 2.1 输入、输出参数选取

由于涡轮出口总温 ( $T50$ )、涡轮出口总压 ( $P50$ )、涡轮出口流量 ( $W50$ ) 是涡轮系统性能的核心指标,能够全面地反映系统的工作状态和健康状况。其中,  $T50$  超出设定的范围,无论是过高、还是过低,都可能对系统的稳定性和性能产生不利影响。因此,监测涡轮系统温度可以提前发现系统是否存在过热或过冷等问题,从而采取适当的措施进行调整或维修;异常的  $P50$  值可能标示着系统存在泄漏、堵塞或其他问题。因此,对涡轮系统压力的监测可以及时发现系统的异常情况,从而减少潜在的故障风险; $W50$  是涡轮系统中控制燃烧过程的关键参数。异常的燃油流量可能表明燃烧过程不稳定或燃料供给存在问题。因此,监测涡轮系统燃油流量可以及时发现系统的故障风险,保障系统的安全稳定运行。在以上分析的基础上,结合发动机测点,最终选择  $T50(^{\circ}R)$ 、 $P50(\text{psia})$ 、 $W50(\text{lbm/s})$  作为涡轮系统数据驱动建模的输出参数。

同时充分考虑航空发动机运行物理原理与监测数据之间的相关性,选定 11 个航空发动机机组运行

参数作为状态预测模型的特征输入,具体如下:高度(alt,单位为ft)、马赫数(Mach)、节气门-分解器角度(TRA,单位为deg)、高压压气机出口总温( $T_{30}$ ,单位为 $^{\circ}\text{R}$ )、高压涡轮出口总温( $T_{48}$ ,单位为 $^{\circ}\text{R}$ )、高压压气机出口净压( $P_{s30}$ ,单位为psia)、物理风扇转速(Nf,单位为r/min)、物理核心速度( $N_c$ ,单位为r/min)、燃油流量(Wf,单位为pps)、高压压气机出口总压( $P_{30}$ ,单位为psia)、高压涡轮出口流量( $W_{48}$ ,单位为pps)。

## 2.2 PCA 降维结果

在PCA处理中,设定了一个阈值,即0.999。这个阈值意味着,当累积贡献率超过0.999时,所选取的主成分已经涵盖了原始数据中99.9%的信息。通过PCA技术,得到了各个主成分的累积贡献率以及相应的变化情况。这些结果在图3中进行了详细展示。

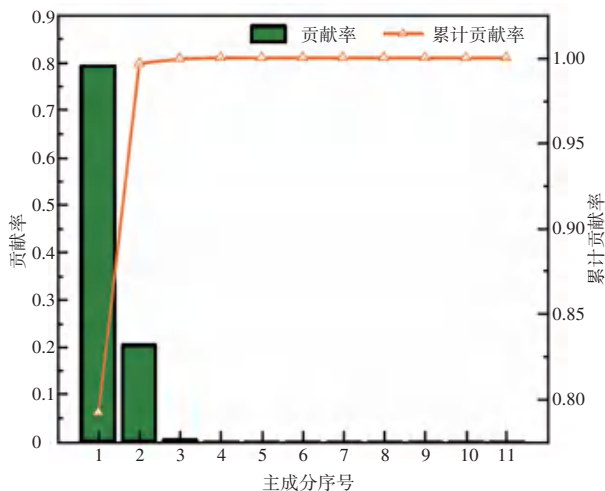


图3 主成分贡献率以及累计贡献率变化

Fig. 3 Variation of principal component contribution rate and cumulative contribution rate

通过分析图3,可以发现,第1主成分对整体的贡献度超过了79%,在所有主成分中占据了较大的比重。当考虑到第3主成分时,整体的贡献度已经累积到了99.9%,而在此之后的主成分对模型的贡献已经变得微乎其微。基于这一结果,选择将前3个主成分作为构建预测模型的关键输入变量,从而将原始的11维数据有效地简化为3维,有效提高了模型的效率和准确性。

在表1中,详细列出了3个主要构成要素的特征值、贡献度以及累积贡献度。首先,通过特征值的计算,得到了每个主构成要素对应的特征向量,进而确定了每个主构成要素的具体数值。

根据表1的数据,当考虑第2个主成分时,特征

值之和已超过1。考虑到累积贡献率和特征值,决定选择前3个主成分作为预测模型的输入参数。这一决策有效地将数据维度从原始的11维降至3维。

表1 前3个主成分信息

Table 1 Information of the top 3 principal components

主成分	特征值	贡献率/%	累计贡献率/%
1	8.714 721	79.225	79.2
2	2.245 686	20.415	99.6
3	0.031 454	0.286	99.9

为了比较使用PCA处理过的数据与未经过处理的数据建模所用时间的不同,选择未经PCA降维的11维数据作为模型的输入,输出仍为3维数据,进行Adam-LSTM数据驱动建模,并与PCA处理后的数据建模时间进行对比,结果见表2。

表2 PCA降维数据与未降维数据建模时间

Table 2 Modeling time for PCA-reduced data and non-reduced data

模型	程序运行时间/s
PCA-Adam-LSTM 数据驱动模型	66.08
Adam-LSTM 数据驱动模型	73.41

仿真结果表明,经PCA处理的数据驱动模型运行时间比未经PCA处理的模型减少了7.33s,缩短了9.99%的时间。

## 2.3 模型性能评估与分析

在进行模型评估时,采用了3种不同的误差指标,以确保评估结果的全面性和准确性。具体来说,选择了平均绝对百分比误差(MAPE)、平均绝对误差(MAE)和均方根误差(RMSE)作为评估指标。这些指标的计算公式如下:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (1)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (3)$$

其中, $n$ 表示数据样本的个数; $\hat{y}_i$ 表示预测值; $y_i$ 表示真实值。

为了评估PCA-Adam-LSTM模型的预测能力,将其与PCA-LSF和PCA-BP模型进行对比。首先,建立了相应的模型。然后,使用测试集来验证模型的预测结果。由于测试样本数据量较大,为了便于展示,从3种输出参数中随机选取了100组数据

进行绘图。 $T50$ 、 $P50$ 、 $W50$  三种参数的预测效果如图 4 所示。

从图 4 的分析结果可以看出,PCA-BP 算法构建的预测模型在预测性能上存在一定的不足,其预测结果与实际值的偏差较大。相比之下,采用

PCA-LSF 和 PCA-Adam-LSTM 算法构建的预测模型在预测效果上更为出色。此外,表 3 详细列出了发动机涡轮系统输出参数的预测评估指标,为进一步分析提供了重要参考。

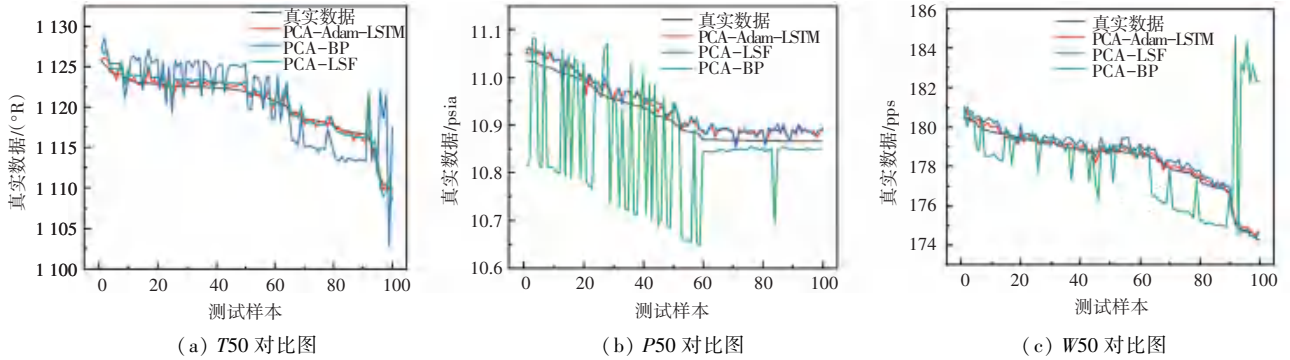


图 4 3 种模型预测值与实际值的对比图

Fig. 4 Comparison of predicted values and actual values for three models

表 3 发动机涡轮输出参数预测评估指标信息

Table 3 Evaluation metrics for turbine output parameter prediction

评估指标		PCA-LSF	PCA-BP	PCA-Adam-LSTM
T50	MAPE	0.111 740	0.351 431	0.088 746
	MAE	1.291 700	0.455 819	1.010 056
	RMSE	1.899 000	0.541 151	1.286 561
P50	MAPE	0.371 311	0.493 449	0.181 728
	MAE	0.044 428	0.656 478	0.021 192
	RMSE	0.058 553	0.926 014	0.024 588
W50	MAPE	0.126 400	1.562 212	0.186 401
	MAE	0.242 281	3.313 574	0.340 636
	RMSE	0.342 038	4.004 179	0.426 939

根据表 3,可以对 3 种模型进行更深入的分析。从  $T50$  预测的角度来看,这 3 个模型的 MAPE 值都小于 0.36,MAE 值都小于 1.30,而 RMSE 值都小于 1.90。在比较其他 2 种方法时,发现 PCA-Adam-LSTM 模型的 MAPE 是最小的,而 PCA-BP 模型的 MAE 和 RMSE 也是最小的;从对  $P50$  预测的角度分析,3 个模型对  $P50$  的 MAPE 均小于 0.50,MAE 均小于 0.66, RMSE 均小于 0.93。对比其它 2 种方法,PCA-Adam-LSTM 模型预测效果最好,MAPE 降低幅度均大于 51.06%,MAE 降低幅度均大于 52.30%,RMSE 降低幅度均大于 57.97%;从对  $W50$  预测的角度分析,3 个模型对  $W50$  的 MAPE 均小于 1.57,MAE 均小于 3.32, RMSE 均小于 0.43。对比其它 2 种方法,PCA-BP 模型的 MAPE、MAE 和 RMSE 预测的预测效果最差。PCA-Adam-LSTM 模型和 PCA-LSF 模型预测效果较好,其中 PCA-LSF

模型的预测效果为最好。综合对  $T50$ 、 $P50$  和  $W50$  的预测效果进行分析,发现 PCA-Adam-LSTM 模型在多个指标下表现出显著优势。对于  $T50$  的预测,3 种模型均表现良好,但在 MAPE 方面,PCA-Adam-LSTM 模型最为出色;而对于  $P50$  的预测,PCA-Adam-LSTM 模型不仅在 MAPE、MAE 和 RMSE 上都优于其他 2 种模型,其预测效果的提升幅度也明显高于其它两者;在  $W50$  的预测中,虽然 3 种模型的表现都可接受,但 PCA-BP 模型表现最差。

总体而言,PCA-Adam-LSTM 模型相比 PCA-BP 模型和 PCA-LSF 模型具有如下明显优势:

(1) 预测准确性更高。尤其在处理  $P50$  的情况下表现突出。

(2) 更好的泛化能力。由于 LSTM 模型的结构特点,PCA-Adam-LSTM 模型能够更好地捕捉数据中的时序信息。

(3) 更好地处理长期依赖关系。对于复杂数据关系的建模,PCA-Adam-LSTM 模型也更为适用。

因此,本文提出的模型融合了 PCA、Adam 和 LSTM 的优点,旨在预测涡轮系统输出参数。该模型不仅缩短了训练时间,并且通过深入分析历史健康数据的内在联系,成功地掌握了其潜在规律。在此基础上,该模型能够对各项参数进行精确的预测,从而为相关领域的研究和应用提供了有力的支持。

### 3 结束语

本文针对传统数据驱动状态预测模型存在输入参数规模过大导致运算负担过大和潜在“鞍点”的

问题,提出了一种结合 PCA、Adam 和 LSTM 的发动机涡轮数据驱动状态预测方法。通过本文的研究,可以得到以下结论:

(1)使用 PCA 降维技术,将原本的 11 维输入参数降至 3 维。在保留超过 99.9%原始数据信息的同时,与未采用 PCA 降维的模型相比,模型优化和训练时间缩短了 2 028.52 s,有效减轻了计算负担。

(2)经过 PCA 降维处理数据,并结合 Adam 优化算法对 LSTM 进行学习率更新,构建了一个用于预测发动机涡轮输出参数的模型。通过与真实数据、其它模型以及 3 个指标的比较,发现基于 PCA-Adam-LSTM 的涡轮数据驱动模型在预测精度和泛化能力方面表现出色,具有显著的实际应用价值。

(3)通过发动机涡轮健康运行的历史数据对  $T50$ 、 $P50$ 、 $W50$  进行预测模型的建立,可以准确预测出发动机涡轮健康运行时的  $T50$ 、 $P50$ 、 $W50$  数值,为发动机涡轮的状态预测提供理论与技术的支持。

## 参考文献

- [1] 李应红,尉询楷,胡金海,等. 航空发动机的智能诊断、建模与预测方法[M]. 北京:科学出版社,2013.
- [2] 殷锴,钟诗胜,那媛,等. 基于 BP 神经网络的航空发动机故障检测技术研究[J]. 航空发动机,2017,43(1):53-57.
- [3] SHEVCHENKO S A, KONOTOPOV O I. The dynamic characteristic of a resonator of the gyroscope based on elastic waves in solids: Finite-element modeling[J]. Physics and Mathematics, 2021,14(2):63-77.
- [4] REBECCA M, MILLER D, GROGAN R. Development of empirical and analytical reaction wheel disturbance models[EB/OL]. (2012-08-22). <https://doi.org/10.2514/6.1999-1204>.
- [5] STETSON J R J B. Reaction wheel low-speed compensation using a dither signal[J]. Journal of Guidance, Control, and Dynamics, 1993,16(4):617-622.
- [6] 周宁宁,卿涛,周刚. 飞轮用轴承组件摩擦力矩特性研究[J]. 空间控制技术与应用,2013,39(3):54-58.
- [7] 国洪立,陈时熠,唐寅,等. 基于 PCA-SSA-LSTM 的余热锅炉状态监测模型[J]. 热力发电,2023,52(10):86-94.
- [8] 范勇,胡名东,杨广栋,等. PCA-GWO-SVR 机器学习用于边坡爆破振动速度峰值预测研究[J]. 振动工程学报,2024,37(8):1431-1441.
- [9] 陈勇,李鹏,张忠军,等. 基于 PCA-GA-LSSVM 的输电线路覆冰负荷在线预测模型[J]. 电力系统保护与控制,2019,47(10):110-119.
- [10] 张厚,刘长良,王梓齐. 基于 PCA 和改进 TOPSIS 法的电厂混煤配比方案综合评价[J]. 热力发电,2019,48(11):73-78.
- [11] 李杨,蓝茂蔚,赵国钦,等. 基于 PCA-PSO-LSSVM 的电站锅炉效率预测模型研究[J]. 热力发电,2021,50(12):43-50.
- [12] 时梦想,胡泓,吴浩,等. 基于 LSTM 神经网络的压电执行器位移迟滞建模[J]. 压电与声光,2023,45(2):231-238.
- [13] 罗广诚,郜家珏,蔡文学. 基于 GRA-LSTM 与 SARIMA 组合模型的季节性时间序列预测[J]. 智能计算机与应用,2021,11(6):195-200.
- [14] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997,9(8):1735-1780.
- [15] 万俊杰,单鸿涛. 基于 WOA 优化 LSTM 神经网络的配电网可靠性评估[J]. 智能计算机与应用,2021,11(10):107-112.
- [16] 纪严杰,樊重俊. 基于 WNR-CLSSA-LSTM 的短期电力负荷预测[J]. 智能计算机与应用,2023,13(7):76-84.
- [17] KUANG Junwei, YANG Hangzhou, LIU Junjiang, et al. Dynamic prediction of cardiovascular disease using improved LSTM[J]. International Journal of Crowd Science, 2019,3(1):14-25.
- [18] WANG Peixiao, WANG Hongen, ZHANG Hengcai, et al. A hybrid Markov and LSTM model for indoor location prediction[J]. IEEE Access, 2019,7:185928-185940.
- [19] QIANG Jiang, TANG Chenglin, CHEN Chen, et al. Stock price forecast based on LSTM neural network[C]//Proceedings of the Twelfth International Conference on Management Science and Engineering Management. Cham:Springer, 2019:393-408.
- [20] 杨斌,杜文娟,王海风. 数据驱动下的虚拟同步发电机等效建模[J]. 电网技术,2020,44(1):35-43.
- [21] 杨威,周义棚. 基于 LSTM 网络预测智能车辆变道的路径规划研究[J]. 智能计算机与应用,2023,13(7):144-149.