

# 基于深度学习的电气大数据故障分析模型研究

蒋丽<sup>1</sup> 崔鸣<sup>2</sup>

(1. 中汽信息科技(天津)有限公司, 天津 300300; 2. 中国汽车战略与政策研究中心, 天津 300300)

【欢迎引用】蒋丽, 崔鸣. 基于深度学习的电气大数据故障分析模型研究[J]. 汽车文摘, 2024(12): 34-39.

【Cite this paper】JIANG L, CUI M. Research on Fault Analysis Model of Electrical Appliance Large Data Based on Deep Learning[J]. Automotive Digest (Chinese), 2024(12): 34-39.

【摘要】汽车智能化技术的快速发展使整车功能日益复杂化、文件开发数激增, 对传统车辆电气故障解析提出新挑战。为了提升电子电气故障分析与诊断能力, 利用智能诊断技术, 构建电气数据分析模型, 辅助人工实现汽车电气系统故障的智能分析与定位。模型构建分为两步: 一是电气测试数据的收集与预处理, 包括报文解析、数据清洗及特征提取; 二是利用深度学习算法构建故障分析模型, 通过模型训练实现智能故障分析与定位。提出基于智能诊断技术的汽车电气数据分析模型, 能够高效解析复杂电气测试数据, 实现故障智能分析与定位, 为电子电气故障排查提供有力支持。

关键词: 大数据分析; 智能辅助诊断; 故障分析建模

中图分类号: TP311.1; U471.1 文献标志码: A DOI: 10.19822/j.cnki.1671-6329.20230115

## Research on Fault Analysis Model of Electrical Appliance Large Data Based on Deep Learning

Jiang Li<sup>1</sup>, Cui Ming<sup>2</sup>

(1. China Auto Information Technology Co., Ltd., Tianjin 300300; 2. China Center for Automotive Strategy and Policy Research, Tianjin 300300)

【Abstract】The rapid development of automotive intelligent technology has led to increased complexity of vehicle functions and a surge in the number of document developments, posing new challenges for the analysis of traditional vehicle electrical faults. In order to enhance the capability of electronic and electrical fault analysis and diagnosis, an intelligent diagnostic technology is utilized to construct an electrical data analysis model, which aids in the intelligent analysis and localization of faults in the automotive electrical system. The model construction is divided into two steps: firstly, the collection and preprocessing of electrical test data, including message parsing, data cleaning, and feature extraction; secondly, the use of deep learning algorithms to build a fault analysis model, achieving intelligent fault analysis and localization through model training. The proposed automotive electrical data analysis model based on intelligent diagnostic technology is capable of efficiently parsing complex electrical test data, realizing intelligent fault analysis and localization, and providing strong support for the troubleshooting of electronic and electrical faults.

Key words: Big data analysis, Intelligent assisted diagnosis, Fault analysis modeling

## 0 引言

随着整车智能化趋势日益显著, 车辆智能驾驶的各项功能不断完善。车辆电子功能的增多和设计逻辑的复杂性提升, 导致故障排查难度增加, 为电气功能的品质提升增加挑战。电气问题的返修除借助原理图、线束图和诊断故障代码(Diagnostic Trouble Code, DTC)外, 还需借助车载网络通信信号、唤醒睡眠机制、

传感器与执行器硬线信号等复杂数据文件进行诊断分析。面对未来整车电气领域技术的不断发展, 电气类数据的未来发展趋势主要体现在以下3个方面:

(1) 整车功能复杂化。随着开发文件不断增多, 电气生准范围及内容不断增多, 整车电气诊断不断复杂化, 车辆电气故障解析的排查点将由2位数攀升至3位数, 传统方式下的解析方式应对艰难。

(2) 电气测试数据持续增加。整车电气测试数据

由MB级增长至GB级甚至TB级,数据解析工作不断复杂化,且车辆故障解析面临更多的电控唤醒机制、标定自学习机制等,对报文解析和逻辑判断技术提出更高的能力要求。

(3)整车总线数据呈几何式增长。整车监听数据由GB级别增长至TB级甚至PB级别。数据解析工作复杂化,类似于高压互锁故障,在借助总线监听、总线开发环境(CAN open environment, CANoe)解析总线故障时,数据收录、转译和分析周期由1个月增加至2个月。

因此,在智能诊断技术背景下研究电气数据分析,构建一套电气数据分析工具,运用数字化技术辅助人工开展数据分析,实现电气系统故障智能分析,可以显著提升故障分析及诊断能力。

在电气大数据故障分析模型的研究中,国内外的研究方法存在很多相似之处。两者均采用了传统统计方法和深度学习方法,并且注重数据的收集、预处理和特征工程等环节。其区别在于国外研究通常具有更大规模的数据集,包含更多样的设备和故障情况,为模型的建立和评估提供了更多信息。且国外的研究更加注重学术研究和算法的发展。国内研究更注重与实际工业应用的结合,对于故障分析模型的实际部署和应用具有更丰富的经验。国内外对电气大数据故障分析模型的研究均取得了一定进展。本文研究成果基于硬件在环仿真(Hardware-In-the-Loop, HIL)台架试验,应用深度学习模型,实时进行故障诊断与预测。着力于实际汽车生产制造过程的电气故障排查,对车辆质量提升具有更高经济价值与社会价值。

## 1 技术架构

本文提出的技术架构主要包括电气数据智能分析与展示、基于深度学习的多目标优化故障识别模型以及基于现有HIL测试台架的故障智能诊断系统3个部分。

(1)电气数据的智能分析和大数据展示旨在提升数据查询和比较的便捷性,为电子电气故障智能诊断、数据智能分析以及电气故障排查提供可视化分析和更加可靠的数据支持。通过对电气设备监测数据处理和分析,实现对电气运行参数、运行状态等数据的实时监控和预报警推送。在系统内注入电气故障信号后,系统基于故障预测模型快速反馈发生故障的电气设备和电气故障的原因,并基于计算的概率进行排序,为业务人员提供维护建议。

(2)基于深度学习的多目标优化故障识别模型旨在提供故障分析的决策优化建议<sup>[1-8]</sup>。通过分析监测到的电气数据,利用故障预测模型预测设备运行状态。针对故障电气设备或元件,智能分析工具根据电子电气系统逻辑、预训练的电子电气系统故障模式模型、故障知识库以及电子电气维修知识库,采用多目标优化技术(如Pareto优化、遗传算法、免疫算法、多梯度下降算法),快速确定故障原因、故障元器件,以及满足目标条件(如维修时间最短、维修成本最低)的最优维修策略建议。

(3)故障智能诊断系统开发旨在实现对现有HIL测试台架数据的无缝连接并进行数据分析<sup>[11-12]</sup>。电子电气故障智能诊断是一个开放的故障预测和智能诊断平台,支持多种数据接入形式。对于在线设备信息接入,该系统能够提供多种标准化接口(Application Programming Interface, API),主要包括REST API、SOAP API以及SQL/No-SQL数据接口。通过数据接口将设备信息快速接入故障智能诊断系统,用于数据分析和故障预测。对于基于文档的离散数据,该系统提供了常用文件格式数据提取,如Excel、CSV、JSON、TXT。针对现有的HIL测试台架数据,该系统可以通过标准化接口直接读取台架测试数据,进行台架数据的智能分析、故障预测以及故障分析和决策。

## 2 技术方法与模型构建

基于智能诊断技术,开发了一套适应整车电子电气系统的电气数据分析工具,可以实现电气故障智能诊断及电气数据的智能分析和大数据展示,便于数据查询和比对。基于深度学习的多目标优化,建立故障预测模型,该模型旨在为故障分析提供决策优化建议。

### 2.1 数据采集融合

多源数据融合对于电气数据分析和数据建模具有重要作用<sup>[13-17]</sup>。存储在不同系统或磁盘中的数据可通过前文提及的多源数据提取组件实现从离散数据到流数据(Streaming Data)格式的转换。流数据是指以实时或连续的方式生成和传输的数据,具有实时性、高速性、多样性以及大规模性的特点。流数据的处理需采用特定的技术和工具,如流处理引擎、复杂事件处理(Complex Event Processing, CEP)系统、实时分析和深度学习算法。本文采用实时分析和机器学习算法对流数据进行实时处理和分析,实现数据的实时洞察和响应功能。数据存储具有分散性,但其存在一定关联性。

为了更有效地分析和处理数据,本研究采用数据融合方法,将分散的数据整合为结构化的表格,便于进行综合分析利用。

数据融合方法的基本原理为将参考参数相同的数据聚合在一起。本项目中,根据使用场景和数据类型,开发了4种多源数据融合组件。

(1)合并数据(Merge Data)。合并来自不同数据源的具有相同参数值的数据,形成一条数据记录,针对不同数据源的数据设置参考参数。

(2)连接数据(Join Data)。根据指定的参数,在参数值一致的情况下,将多个数据源的数据扩展至现有数据记录中,使用Join Using参数指定参考参数。

(3)分组数据(Group Data)。将具有相同参数值的所有数据记录分组到单一数据记录中,需要指定参考参数,使用Group Using指定分组操作。

(4)透视数据(Pivot Data)。透视数据可以将狭窄的数据转换为宽数据表。根据指定的列,将列的值转换成列名,将另一列的值转换为列值,并使用键属性值进行合并。

## 2.2 特征参数与建模

由于数据来源和格式具有多样性,在接入电气故障诊断系统时,首先需基于数据结构和特点采取数据预处理步骤,如数据合并、数据拆分和数据转化,将相关数据聚合成一个标准化、结构化的表格形式。完成数据聚合后,需对整合后的数据进行进一步处理,对其异常数据和无效数据等进行清洗、填充或者替换处理,提升数据质量。获得高质量的电气数据后,使用数据分析模块对数据进行初步的探索分析(如聚类分析、统计分析),了解数据的分布情况以及分布趋势,辅助确定特征参数。可选数据获取途径包括通过汽车车载自动诊断系统(On-Board Diagnostics, OBD)接口获取CAN总线上的数据,或基于HIL台架获取一段时间内电气系统的运行数据。

本研究采用基于深度学习的检测模型构建汽车电气系统运行状态特征。该模型采用深度学习框架下的自编码器,对输入数据进行降维,得到表征汽车电气系统运行状态的特征。由于上述获取的运行数据类型较多且汽车各模块耦合运行,不同类型的运行数据间存在强相互联系,因此运行数据中保留的噪声、细节和重复信息较多。本研究采用自编码对运行数据进行编码,可以减少数据类型。为了明确区分和描述数据处理过程,将自编码器的输入定义为“数据”,输出定义为“特征”,输出特征的维度低于输入数据的数据类型,实

现了数据的降维。

## 2.3 模型训练与优化

### 2.3.1 构建正常运行数据样本库

正常检测模型的特性在于其能够识别正常运行数据。若输入该模型的运行数据为正常运行数据,其输出特征与正常特征库中的特征相匹配。基于该特性,构建正常运行数据样本库,对待训练模型进行训练,从而优化待训练模型的编码器参数,使其自编码器满足上述特性,成为正常检测模型。

### 2.3.2 模型训练

在对待训练模型进行训练过程中,将第一组正常运行数据样本和第二组正常运行数据样本先后输入至所述待训练模型,通过最小化所述第一组正常运行数据样本与编解码后数据的距离以及所述第二组正常运行数据样本的编码后特征与所述第一组正常运行数据样本的编码后特征的距离,优化所述待训练模型的编码器参数。解码器的参数采用预设参数,或在每次训练中自编码器一起进行优化。

训练过程中构建损失函数,用于训练正常检测模型:

$$L = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 + a \times \frac{1}{d} \sum_{i=1}^d (z'_i - z_i)^2 \quad (1)$$

式中: $L$ 为损失量, $a$ 为损失偏置, $i$ 为第 $i$ 次量, $n$ 为一组正常运行数据样本中数据类型的总数, $x_i$ 为第一组正常运行数据样本的编解码后数据, $y_i$ 为第一组正常运行数据样本, $d$ 为一组正常运行数据样本的编码后特征的维度总数, $z_i$ 为第一组运行数据样本的编码后特征, $z'_i$ 为第二组正常行数据样本的编码后特征,表示正常运行对应的损失权重。

在具体实施过程中,根据所述多个输出特征构建正常特征库。从所述多个输出特征中,选取相关性最强的输出特征(至少1个)构成正常特征库。在特征比对过程中,采用相关性度量标准判断特征匹配情况。若一特征与正常特征库中的任一特征的相关性足够高(例如欧式距离 $<0.01$ ),则判定该特征与正常特征库中的特征匹配。在模型应用中,若输入正常检测模型的正常运行数据样本为正常运行一段时间产生的运行数据,则正常检测模型输出的特征为多条曲线或二维矩阵。每条曲线反映输出特征中的一个维度随时间的变化规律,如当 $d=4$ 时,存在4条曲线,维矩阵的行和列分别为输出特征的不同维度和不同时刻。正常特征库中的特征也可以是多条曲线或二维矩阵。在特征比对过程中,若某特征中包括的每条曲线与正常特征库中

任一特征的每条曲线均匹配,则该特征与正常特征库中的特征匹配;若某特征对应的二维矩阵与正常特征库中任一特征对应的二维矩阵中每个元素均匹配,则该特征与正常特征库中的特征匹配。

## 2.4 数据识别

根据特征参数构建预测模型和故障分析及决策优化的对训练数据质量和数量的要求,系统支持使用内置或开发的标准模块构建台架试验测试用例 workflow,基于设定范围自动化生成对应的 HIL/试车测试工况场景<sup>[18-23]</sup>。具体故障数据识别流程如图 1 所示。

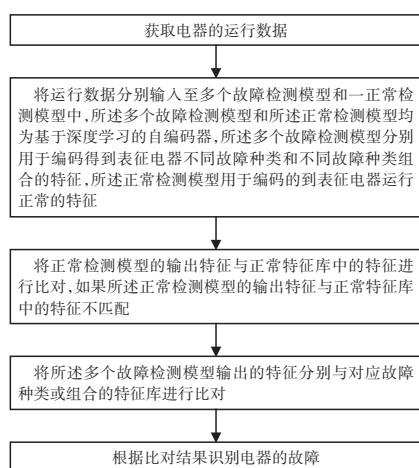


图1 故障数据识别流程

本研究采用了3类结构相同但参数不同的自编码器,每个自编码器对应一种运行状态。正常检测模型的特性为若输入该模型的运行数据为正常运行数据,则该模型输出的特征与正常特征库中的特征匹配。故障种类的故障检测模型的特性为若输入该模型的运行数据为出现该种故障时的运行数据,则该模型输出的特征与该故障种类对应的特征库中的特征匹配。故障种类组合对应的故障检测模型的特性是若输入该模型的运行数据为出现该故障种类组合时的运行数据,则该模型输出的特征与该故障种类组合对应的特征库中的特征匹配。

首次使用所述正常检测模型的输出特征与正常特征库进行比对,若所述正常检测模型的输出特征与正常特征库中的特征不匹配,表明电气发生了故障。为了确定具体故障种类或组合,分别将每个故障检测模型输出的特征与对应的特征库进行比对。

本文中所开发的电子电气故障智能诊断系统包含多数据源数据接入、数据分析、数据清洗、数据模型开发、多参数优化以及数据展示等模块。数据提取工具集通过数据接口实现在线数据和离线数据的接入,实时基于监测电气数据进行预测未来1小时或3小时内

电气设备发生故障的可能性。针对可能发生故障或已发生故障信息,故障分析模型基于进化算法、Pareto多参数优化算法以及故障知识库快速反馈引起故障的原因和故障电气元器件。

## 3 故障智能诊断系统实现

汽车电控域故障智能诊断系统采用目前比较成熟的统计分析、深度学习等大数据挖掘技术构建电气系统电气系统数据智能分析、故障预测、故障智能诊断分析模型,具体电气故障预测模型如图2所示。根据监测到的电气数据,利用故障预测模型预测设备运行状态。对于故障电气设备或元件,故障智能分析工具及根据电子电气系统逻辑,预训练的电子电气系统故障模式模型、故障知识库、电子电气维修知识库等,采用多目标优化技术(如 Pareto 优化、遗传算法、免疫算法、PSO、多梯度下降算法),快速反馈故障原因、故障元器件,以及满足目标条件(如维修时间最短、维修成本最低)的最优维修策略建议。实现电气系统数据在线智能分析,动态展示设备运行参数、运行状态、故障发生概率以及故障发生原因和故障电气元件指认,缩短电气系统监听数据收录、转译和分析周期,提高故障维护效率,降低安全事故发生概率。

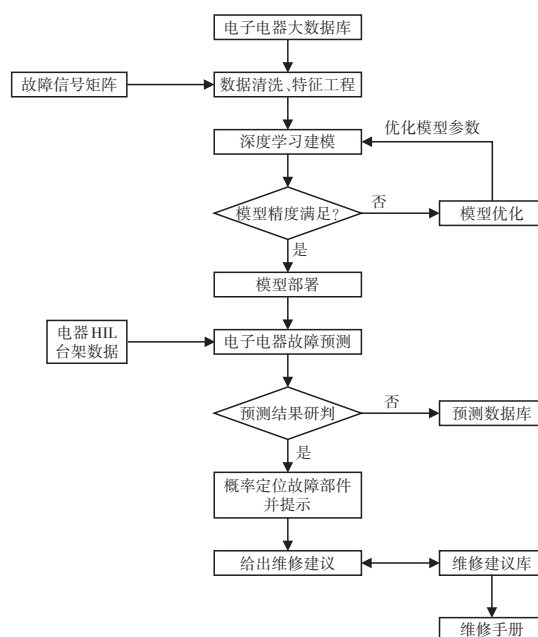


图2 电气故障预测模型

(1)无缝连接 HIL 测试台架数据并设计对应的 HIL/试车测试工况场景。电控域故障智能诊断是一个开放的故障预测和智能诊断平台,平台支持多种数据接入形式。对于在线设备信息接入,能够提供多种标准化 API 接口,如 REST API、SOAP API 以及 SQL/

No-SQL数据接口。通过数据接口将设备信息快速接入故障智能诊断系统,用于数据分析和故障预测,如图3所示,可在平台中选择车型及控制器,导入台架数据。对于基于文档的离散数据,该系统提供了常用文件格式数据提取,如Excel、CSV、JSON、TXT等。针对现有的HIL测试台架数据,该系统可以通过标准化接口直接读取台架测试数据,进行台架数据的智能分析、故障预测以及故障分析和决策。根据系统构建预测模型和故障分析及决策优化的对训练数据质量和数量的需要,系统支持使用内置或开发的标准模块构建台架试验测试用例试验设计(Design Of Experiment) workflow,基于设

定范围自动化生成对应的HIL/试车测试工况场景。

(2)故障信息采集、数学模型分析、分析结果智能呈现<sup>[24-25]</sup>。如图3所示,数据提取模块通过数据接口实现在线数据和离线数据的接入与查询,包括故障数据和待识别数据。实时基于监测电气数据预测电气设备发生故障的可能性。针对可能发生的故障,或者已经发生的故障信息,故障分析模型可以基于多参数优化算法和故障知识库快速反馈故障原因。此外,基于监测的电气数据和基于模型预测的电气故障信息,使用数据可视化模块中可预定义的各种图表进行智能展示,如热力图、趋势图,如图4所示。

车型	故障代码	代码描述	ECU	故障现象	维修措施
C100	U010387	电子换挡器节点丢失	ACM	无法启动	确认库存没有老版本软件的零件,供应商发货为新版本软件
C100	U040286	变速器控制单元报文内容故障	ACM	ACM写入VIN成功,但读取为0	经过LOG确认,读出来的VIN不为0,且和写入的一致后续持续跟踪,再出现后进行排查
C100	U041586	ESP报文内容故障	ACM	ACM:U041586-ESP报文内容故障	修改软件
C100	B000111	驾驶员前气囊-对地短路	ACU	ACU后排右侧对地短路B003811	工厂排查线束连接
C100	B00011B	驾驶员前气囊-电阻值过高	ACU	ACU VIN 信息错误	重新进行KOEO
C100	B001056	前排乘客前气囊-配置错误	ACU	安全气囊报警灯常亮	EOL配置码写入,已完成写入
C100	B100016	BCM报故障多节点丢失	BCM	新软件已经体现2142、配置未体现,二者不匹配导致	重新匹都车型配置码
C100	B102D12	单体电压过低故障	BCM	BCM DTC:B102D12后视镜倒档输出故障	软件更新
C100	U014687	总电压轻微过高报警	BCM	BCM DTC:U016987天窗控制器,节点丢失	BCM DTC:U016987天窗控制器盖板材料确认接插件是否松动

图3 故障信息采集查询

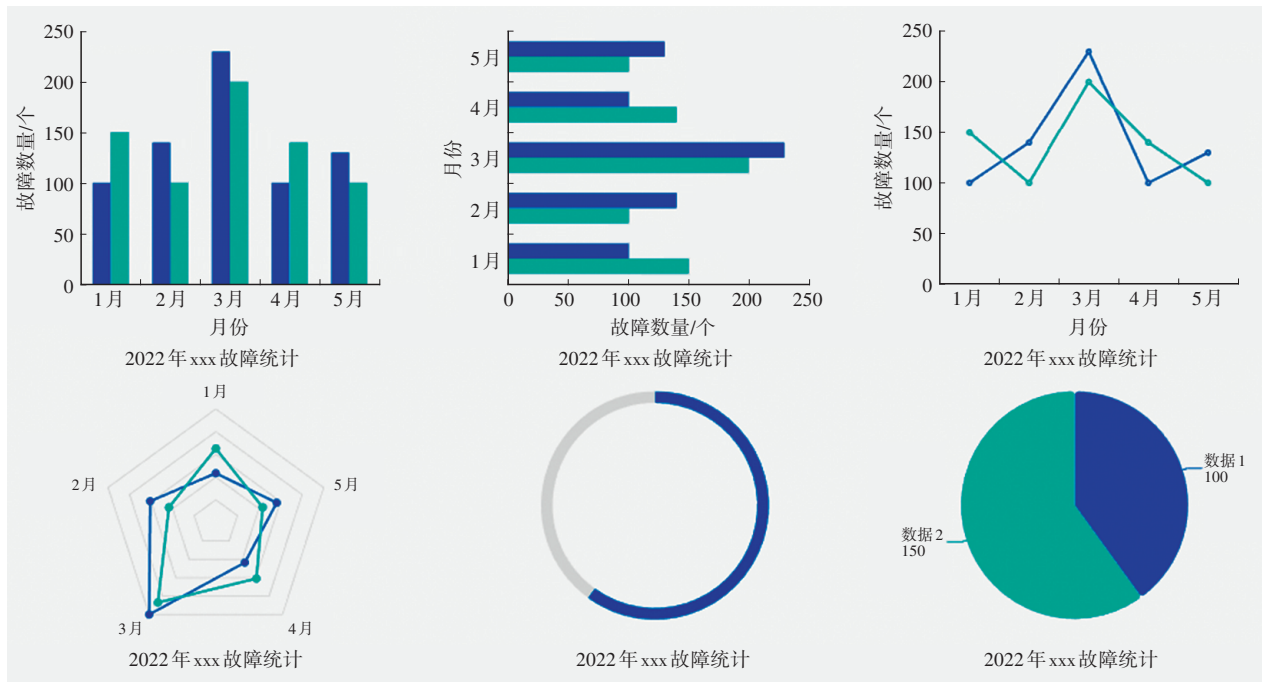


图4 图表智能展示

#### 4 结束语

本项目成功设计并开发了电气故障诊断智能化分析软件系统,该系统基于大数据分析技术,实现了对接入系统电气数据的全面管理与高效利用。系统能够实

时展示数据变化、电气设备运行状态,精确捕捉发生故障的时间点,并智能分析故障原因。通过对出现故障的电气设备进行深度剖析,系统能够导出所有故障信息,并有效识别出引发故障最为频繁的电气设备及其主要原因。该工具充分利用数字化技术,极大地提升

了人工数据分析工作效率,实现了电气系统故障的智能分析,显著提升了电子电气领域的故障分析及诊断能力,为汽车电气故障排查与维护提供了强有力的技术支持。

#### 参 考 文 献

- [1] 杨庆川. 基于深度学习的发动机故障检测[D]. 杭州: 杭州电子科技大学, 2021.
- [2] 单凯旋. 基于深度学习的发动机故障预测方法研究[D]. 天津: 中国民航大学, 2019.
- [3] 梅本祥. 基于MVC模式的车载服务平台监控管理系统研究[D]. 广州: 广东工业大学, 2018.
- [4] 余志生. 汽车理论[M]. 5版. 北京: 机械工业出版社, 2009.
- [5] 胡天杰. 基于模型的ECU硬件在环仿真研究[D]. 昆明: 昆明理工大学, 2020.
- [6] 王望予. 汽车设计[M]. 4版. 北京: 机械工业出版社, 2004.
- [7] 张文杰, 李长龙, 洪宇, 等. 基于大数据技术的远程诊断系统在汽车上的应用[J]. 汽车文摘, 2022(4): 30-33.
- [8] 张佳琪. 基于Java Web的车辆试验管理信息系统的设计与实现[D]. 北京: 北京交通大学, 2016.
- [9] 卢绍兵. 基于Python的混合语言编程及其实现研究[J]. 科技资讯, 2022, 20(14): 31-33.
- [10] 张巍. 基于Python的车辆检测算法的实现[J]. 机械管理开发, 2019, 34(12): 258-261.
- [11] 程学旗, 靳小龙, 王元卓, 等. 大数据系统和分析技术综述[J]. 软件学报, 2014, 25(9): 1889-1908.
- [12] 赵力. 基于大数据处理的车辆监控系统的设计与实现[D]. 北京: 北京交通大学, 2019.
- [13] 曾宪宇. 基于大数据技术的车辆监控系统的优化与实现[D]. 长春: 吉林大学, 2016.
- [14] 李国杰, 程学旗. 大数据研究: 未来科技及经济社会发展的重大战略领域——大数据的研究现状与科学思考[J]. 中国科学院院刊, 2012, 27(6): 647-657.
- [15] 曾雷. 大数据研究综述[J]. 软件导刊, 2015, 14(8): 1-2.
- [16] 付岩, 李洪茹. 关于整车企业数字化转型系统架构的研究[J]. 汽车文摘, 2022(7): 20-26.
- [17] 王茹葳. Java编程语言在大数据开发中的应用[J]. 电子技术, 2022, 51(1): 160-161.
- [18] 郭阳, 常英贤. 浅谈Java语言在计算机软件开发中的应用[J]. 数字通信世界, 2022(1): 88-90+94.
- [19] 张佳琪. 基于Java Web的车辆试验管理信息系统的设计与实现[D]. 北京: 北京交通大学, 2016.
- [20] 李健. 基于深度学习的变循环发动机气路故障诊断[D]. 上海: 上海交通大学, 2019.
- [21] 洪骥宇. 基于深度学习的航空发动机可靠性分析[D]. 南京: 南京航空航天大学, 2018.
- [22] 周亦人, 邱小林, 郭志强. 基于深度学习理论的发动机气门机构故障识别[J]. 制造业自动化, 2017, 39(11): 89-93.
- [23] 严冰. 深度学习在航空发动机故障诊断中的应用[D]. 上海: 上海交通大学, 2017.
- [24] OSTROWSKI K, BIRMAN K, DOLEV D. Extensible Architecture for High-Performance, Scalable, Reliable Publish-Subscribe Eventing and Notification[J]. International Journal of Web Services Research, 2007, 4(4): 18-58.
- [25] YANG J. Web Service Componentization[J]. Communications of the ACM, 2003, 46(10): 35-40.

(责任编辑 梵玲)