

汽车动力智能故障诊断模型构建及检验

张性伟, 谢振, 张赛文, 代云川, 海哈小东, 果机布旦

四川交通职业技术学院, 四川 成都 611130

DOI:10.61369/UAID.2025020020

摘 要 : 汽车动力系统故障诊断是保障行车安全、提升运维效率的重中之重, 本文聚焦智能故障诊断模型的理论构建及验证, 探讨基于深度学习的多模态信号融合分析框架。通过构建包含信号感知层、特征融合层、智能诊断层及决策输出层的理论模型, 并采用蒙特卡洛交叉验证与迁移学习理论模拟, 证明了其在噪声干扰与工况迁移下的强泛化性与鲁棒性, 以期汽车动力系统智能化运维提供了坚实的理论支撑与方法论指导。

关 键 词 : 动力系统; 故障诊断; 深度学习; 特征融合; 图卷积网络

Construction and Verification of Intelligent Fault Diagnosis Model for Automotive Power

Zhang Xingwei, Xie Zhen, Zhang Saiwen, Dai Yunchuan, Haiha Xiaodong, Guoji Buqie

Sichuan Communications College, Chengdu, Sichuan 611130

Abstract : Fault diagnosis of automotive power systems is of Paramount importance for ensuring driving safety and enhancing operation and maintenance efficiency. This paper focuses on the theoretical construction and verification of intelligent fault diagnosis models, and explores a multimodal signal fusion analysis framework based on deep learning. By constructing a theoretical model that includes a signal perception layer, a feature fusion layer, an intelligent diagnosis layer and a decision output layer, and using Monte Carlo cross-validation and transfer learning theory simulation, its strong generalization and robustness under noise interference and working condition transfer have been proved, with the aim of providing a solid theoretical support and methodological guidance for the intelligent operation and maintenance of automotive power systems.

Keywords : power system; fault diagnosis; deep learning; feature fusion; graph convolutional network

引言

汽车动力系统作为整车的“心脏”, 其运行状态直接关乎车辆的动力性、经济性与安全性, 因此汽车动力系统故障的诊断显得极其重要, 而传统故障诊断方法主要依赖阈值判断、专家经验或浅层信号分析, 在应对复杂工况、微弱故障及非线性系统时存在诊断精度低、泛化能力弱、依赖先验知识等固有局限。近年来以深度学习为代表的人工智能技术凭借其强大的特征自主提取与复杂模式识别能力, 为动力系统智能故障诊断开辟了新路径。本文旨在构建一种融合多源异构数据、具备强抗噪能力与工况适应性的智能故障诊断理论模型, 并深入探讨其内在机理与验证方法, 研究将严格遵循理论分析范式, 不涉及具体实验数据表格, 着重从信息处理流程、模型结构设计、算法交互机制、性能验证理论等维度进行深度剖析, 力求为汽车动力系统智能诊断提供系统化的理论框架与创新思路。

一、智能故障诊断模型的理论构建

本文提出的汽车动力智能故障诊断模型是一个层次化、信息流驱动的理论架构, 其核心在于实现多源异构数据的高效融合与深层故障模式的精准识别, 模型由四个相互关联的理论层级构成。

(一) 信号感知与预处理层

信号感知与预处理层也是模型的信息输入端, 负责接收动力系统运行过程中产生的多模态传感器信号, 预处理层的目标是提

升信号的信噪比, 确保后续特征提取的有效性, 同时为多源信号的时空对齐提供理论基础。典型信号源主要包括振动信号——反映发动机、变速箱等旋转机械部件的机械状态, 其预处理核心在于降噪与特征增强, 理论方法涉及自适应小波阈值去噪、经验模态分解及其改进算法以应对信号非平稳特性, 以及基于奇异值分解的信号增强技术; 声学信号——蕴含丰富的机械摩擦、气流脉动、燃烧状态信息, 其处理难点在于环境噪声抑制与目标声源分离, 理论支撑包括盲源分离、基于麦克风阵列的波束形成技术理论; 缸压信号——直接反映发动机燃烧过程状态, 需进行精确的

循环事件同步与压力漂移校正；控制器局域网总线数据流——包含丰富的控制器状态参数与执行器反馈，其预处理侧重于异常值检测与信号同步对齐。

（二）深度特征提取与融合层

深度特征提取与融合层是模型智能化的核心，负责从预处理后的原始信号中自动学习并融合具有高度判别性的故障特征。这一层的深度特征提取主要包括面向振动/声学信号和面向时序依赖性强的信号：深度卷积神经网络通过其卷积层、池化层的堆叠，自动学习信号中的局部时频模式与空间结构特征，其理论优势在于局部感知、权重共享、空间下采样带来的平移不变性与特征抽象能力，针对信号的非平稳特性，结合连续小波变换或短时傅里叶变换将一维信号转化为时频图像，再利用深度卷积神经网络进行图像特征提取，是理论上的有效途径；长短期记忆网络及其变体利用其内部的门控机制比如遗忘门、输入门、输出门等，建模长距离时间依赖关系，有效捕捉故障在时间维度上的演化模式，深度置信网络通过多层受限玻尔兹曼机的堆叠与逐层预训练，可有效提取数据的高阶统计特征^[1]。

本文模型采用基于注意力机制（Attention Mechanism）的深度融合理论，为来自不同模态或不同特征提取器的特征向量分配动态权重，该权重由特征自身的重要性及其与当前诊断任务的关联性共同决定，通过一个可学习的查询-键值映射函数实现。模型能自适应聚焦于与当前故障最相关的特征通道、抑制冗余或噪声干扰大的特征，实现“特征层面的智能选择”，显著提升融合特征的判别力与鲁棒性。

（三）智能诊断层

智能诊断层通过接收融合后的高维特征向量，执行最终的故障模式识别与分类任务。为提升对复杂系统拓扑关系与全局上下文的理解能力，本文创新性地融合两种理论模型。其一是图卷积网络，动力系统可抽象为一个图结构，节点代表关键部件，边代表部件间的物理连接、能量传递或功能依赖关系；图卷积网络通过在图的节点与边上定义卷积操作，能够有效聚合邻居节点信息，捕捉系统固有的拓扑结构信息，识别由部件间相互作用引发的复合故障或传播性故障。

其二是改进 Transformer 模型，标准 Transformer 基于自注意力机制，擅长捕捉长距离依赖与全局上下文信息，针对故障诊断任务对其进行理论改进：引入相对位置编码，更贴合传感器信号在时间或空间上的相对位置关系；设计面向故障诊断的预训练任务，利用海量无标签数据进行模型初始化，提升小样本故障诊断能力^[2]。将图卷积网络学习到的拓扑结构信息作为先验知识嵌入到 Transformer 的自注意力计算中，或利用 Transformer 处理图卷积网络输出的节点特征序列，实现图结构信息与全局上下文信息的深度耦合，最终输出层采用 Softmax 函数，得到各预设故障模式的发生概率分布。

（四）决策输出与解释层

决策输出与解释层是将智能诊断层的概率输出转化为可理解的诊断结论与辅助决策信息。一方面基于最大后验概率准则或设定诊断阈值，确定最终的故障类型及置信度，另一方面为增强模

型可信度与可接受度，集成事后解释方法。比如 LIME 通过局部拟合一个可解释模型来近似复杂模型在特定样本附近的决策行为；SHAP 则基于合作博弈理论，计算每个输入特征对最终诊断结果的贡献度^[3]。这些理论方法能有效标识出对特定故障诊断起关键作用的原始信号片段或特征维度，为维修人员提供故障定位与成因分析的线索。

二、模型检验的理论框架与方法

为确保所构建智能诊断模型的可靠性与实用性，需建立严格的理论检验框架，重点评估其泛化性、鲁棒性与不确定性量化能力。

（一）泛化性能检验理论

本文采用蒙特卡洛交叉验证在理论上模拟数据划分过程，通过多次随机将有限的理论数据划分为训练集、验证集和测试集如 70%/15%/15% 等等，重复进行模型训练与测试，最终的泛化性能指标如平均准确率、精确率、召回率并取多次测试结果的平均值与方差，该方法理论上能更可靠地估计模型在未知数据上的期望性能，减少单次划分的随机性影响。此外还通过迁移学习理论验证模拟模型在工况迁移下的适应能力，检验方法包括：通过在损失函数中引入域差异度量或对抗性训练，迫使模型学习域不变特征，检验模型在目标域上的诊断性能；利用元学习如 Model-Agnostic Meta-Learning, MAML 或基于度量的方法如 Prototypical Networks^[4]，训练模型具备快速适应新故障类型的能力，检验其在少量目标域样本支持下的诊断准确率。

（二）鲁棒性检验理论

评估模型在噪声干扰、传感器失效、数据缺失等非理想条件下的稳定性。其一是尚缺少理论分析模型在不同信噪比（SNR）的加性高斯白噪声、脉冲噪声或实际采集环境噪声干扰下的性能衰减曲线，探究模型各层（特别是特征融合层）对噪声的抑制能力。其二是尚缺少在理论上模拟部分传感器通道完全失效或随机丢失大量数据点的情况，检验模型在特征缺失条件下的诊断能力，评估特征融合策略（如注意力机制）是否能有效降低对单一信号源的依赖，或利用数据插补理论（如矩阵补全、基于生成对抗网络的数据生成）进行补救。

（三）不确定性量化理论

智能诊断模型不仅应给出诊断结果，还应评估该结果的可靠性。一方面智能诊断模型通过贝叶斯深度学习，将模型权重视为随机变量，引入先验分布，通过变分推断或蒙特卡洛方法，如 Dropout as a Bayesian Approximation 近似后验分布，模型预测时可通过多次前向传播采样获得预测分布，计算预测熵或互信息作为不确定性估计。另一方通过全面集成学习训练多个结构不同或初始化不同的基诊断模型，如不同架构的 CNN-LSTM、不同的 GCN 层数、不同的 Transformer 头数等等；通过集成其预测结果（如平均、加权投票），并以预测结果的方差或分歧度作为不确定性度量，高不确定性提示诊断结果可能不可靠，需人工介入核查。

三、讨论与理论挑战

本文构建的智能故障诊断模型理论框架，通过深度特征融合、图结构建模、改进 Transformer 应用及严格的验证理论，展现出应对汽车动力系统复杂故障诊断任务的潜力，然而其走向成熟应用仍面临诸多理论挑战。

首先是跨车型 / 跨平台的泛化适用性^[5]，当前模型高度依赖训练数据所反映的特定系统特性，如何设计更具普适性的特征表示与模型架构，实现模型在不同类型发动机、变速箱甚至不同品牌车型间的有效迁移，是理论研究的重点，元学习、自监督预训练、通用系统建模语言是可能方向。其次是零样本 / 少样本故障诊断^[6]，对于训练数据中从未出现过的全新故障模式（零样本），或仅有极少数样本的罕见故障（少样本），模型识别能力有限，亟须发展结合因果推理、知识图谱（融入物理模型、故障树等先验知识）与生成模型（合成故障数据）的理论方法。再次是边缘计算部署理论，模型复杂度与实时性、车载嵌入式系统资源有限性之间存在矛盾，模型轻量化理论（如剪枝、量化、知识蒸馏）、硬件感知神经网络架构搜索（NAS）及高效推理算法是实现在线诊断的关键。第四是可解释性与可信性，现有 XAI 方法（如 LIME, SHAP）提供的解释往往是局部且事后的，发展具有内在

可解释性的模型结构（如原型网络、可解释规则嵌入）以及能反映系统物理机制的解释，对于获得维修人员与监管机构的信任至关重要。第五是多目标优化理论，诊断模型需同时优化准确性、实时性、鲁棒性、资源消耗等多个目标，如何在这些相互冲突的目标间进行有效权衡，需要更先进的多目标优化理论与评估指标体系。

结论：本文系统构建了一种面向汽车动力系统的智能故障诊断理论模型，并深入阐述了其内在机制与检验方法，在模型检验方面建立了基于蒙特卡洛交叉验证、迁移学习理论的泛化性评估框架，系统分析了模型的噪声鲁棒性与不确定性量化能力。研究表明所提出的理论框架为克服传统诊断方法的局限性、实现汽车动力系统状态监测与故障诊断的智能化升级提供了可行的技术路径。其核心价值在于实现了从数据到知识、再到智能决策的自动化闭环，显著提升了诊断的精度、效率与可靠性，然而模型的全面工程化应用仍需在跨平台泛化、小样本学习、模型轻量化、内在可解释性等前沿理论方向取得突破。未来研究应更紧密地结合物理机理与数据驱动，发展融合先验知识的可解释、可信、强适应性的新一代智能诊断理论，并积极探索其在边缘计算环境下的高效部署方案，为汽车工业的智能化运维与安全保障提供持续的理论创新动力。

参考文献

- [1] 杨明, 丁能根, 郑四发, 等. 汽车发动机故障诊断的深度学习 [J]. 清华大学学报 (自然科学版), 2017, 57(11): 1183-1189.
- [2] Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019.
- [3] Lundberg, S.M., & Lee, S.I. A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems, 2017.
- [4] Snell, J., Swersky, K., & Zemel, R. Prototypical Networks for Few-shot Learning. Advances in Neural Information Processing Systems, 2017.
- [5] 王飞跃, 曾大军, 李众. 平行智能: 复杂系统的计算化建模、网络化控制与智能化决策 [J]. 自动化学报, 2015, 41(4): 605-612.