

基于数据驱动的航空发动机故障诊断

Fault Diagnosis of Aero Engine Based on Data-Driven

孙泽茹 / 中国航空发动机研究院

航空发动机故障诊断技术是实现系统视情维修、降低飞行成本、确保系统可靠性的关键，也是航空发动机健康管理的重要内容。基于数据驱动的航空发动机故障诊断依据发动机测量参数的变化来判断和识别发动机的工作状态和故障，对故障进行早期预报，进而保证发动机安全、稳定、可靠地工作，目前已发展成为航空发动机故障诊断领域广泛应用的方法。

航空发动机是在高温、高压、高速、高强度和变负荷等极端条件下工作的复杂大型装备，对飞机的安全性和可靠性有极为重要的影响，航空发动机能否健康、稳定地工作直接影响飞机的飞行安全。但其工作环境复杂恶劣，容易导致故障发生，轻则引起发动机性能下降，致使飞机迫降停飞；重则会损坏发动机，诱发飞行事故、危及飞行安全。为进一步提高航空发动机的安全性，急需开展航空发动机预测与健康管理的研 究，而故障诊断是其中最为关键和重要的内容。在飞机的常规维修及诊断过程中，维修费用巨大，有必要采用智能方法对航空发动机进行故障诊断和预测，既节省检修维护费用，又能快速、准确地确定故障的部位，保障飞行安全。

为了降低航空发动机的维修成本，使发动机高效运转，利用发动机运行的状态数据进行故障诊断的研究越来越受到重视。目前，基于数据驱动的航空发动机故障诊断方法已渐为成熟。

基于数据驱动的航空发动机故障诊断方法概述

基于物理模型的方法需要了解设备

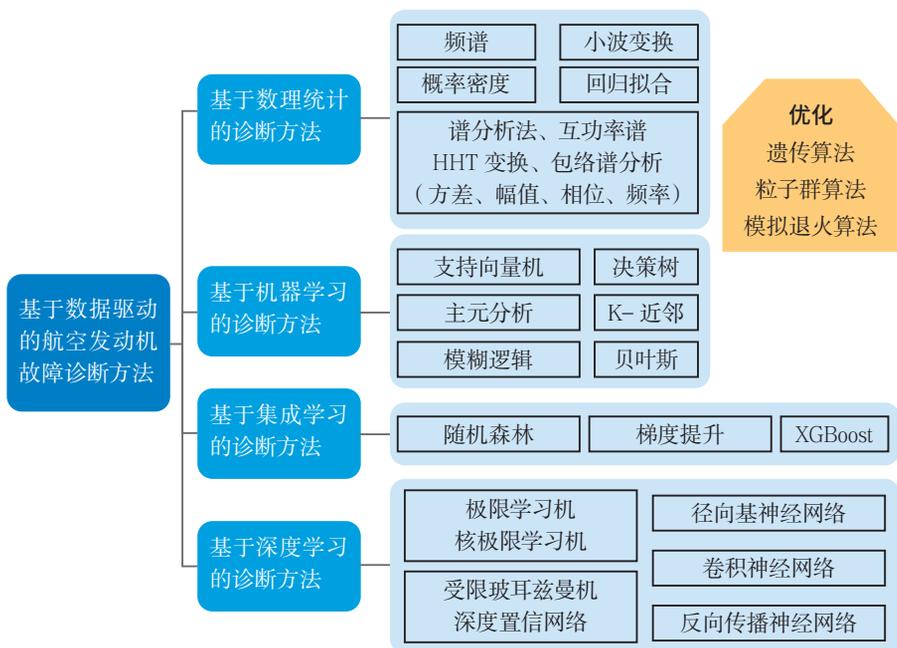


图1 基于数据驱动的航空发动机故障诊断方法

的运行特点、结构功能特征等信息，依据这些信息建立数学模型对设备进行故障诊断。但是该类方法非常依赖模型的精度，随着发动机的结构越来越复杂，发动机各部件之间故障存在耦合，基于物理模型的诊断方法难以对发动机进行准确建模，具有较大的局限性。与此同时，发动机的监控数据不断累积，给基于数据驱动的方法提供了海量数据基础。航空发动机在运行过程中，伴

随故障的发生，必然产生诸如振动、噪声、温度和压力等物理参数的变化。基于数据驱动的航空发动机故障诊断依据这些参数的变化判断和识别航空发动机的工作状态，对故障进行早期预报、识别，做到防患于未然。对航空发动机故障诊断与预测领域的国内外相关文献进行分析，现有基于数据驱动的航空发动机故障诊断方法（见图1）主要包括基于数理统计的诊断方法、基于机器学习的

诊断方法、基于集成学习的诊断方法和基于深度学习的诊断方法。

基于数理统计的故障诊断

基于数理统计的诊断方法通过对过程数据统计量的分析,利用频谱、小波变换、概率密度、回归拟合、相关函数等数据处理模型,直接提取测量参数包括时域、频域、时频域等的特征,进而可以从其变化中提取特征,针对特定的变量设定特定的门限值来检测异常。基于数理统计的故障诊断方法故障判据较为清晰、故障判断准则建立相对容易。

时域特征包括有量纲特征如一阶矩均值、幅值、均方根、二阶中心矩方差等和无量纲(量纲一)特征如波形因子、脉冲因子、裕度、峭度、三阶中心矩偏度、四阶中心矩峰度等。能代表发动机故障发生前后参数变化的特征是发动机故障诊断的关键,可提取测量参数偏差的阈值特征,通过对比历史数据获取表征故障发生后测量参数变化作为监测值。范满意等利用发动机基准模型获取测量参数的偏差,将发动机测量参数偏差的统计分布特性作为故障特征,基于D-S证据理论进行融合故障诊断提高了准确率^[1]。

频域特征提取主要包括频带能量提取及特征频率提取。频带能量提取是指在频带内指定的频段内提取所对应的能量,特征频率提取是指在特定的频率点提取该点的幅值。特征频率的分析方法主要有傅里叶变换、边际谱、包络谱和希尔伯特-黄变换(HHT)等。傅里叶变换主要处理平稳信号;边际谱可以处理非平稳信号,准确反映信号的实际频率成分;包络谱剔除了不必要的频率

干扰,对冲击事件的故障比较敏感,更能够凸显故障特征频率;HHT对信号经验模态分解(EMD)后的各个分量进行希尔伯特变换,能够得到有实际意义的瞬时频率。

常用的时频域分析提取方法包括短时傅里叶变换(STFT)和小波分析,不同于短时傅里叶变换,小波分析所采用的基函数是一种幅值衰减、可伸缩、可平移的小波基,而短时傅里叶变换基于快速傅里叶变换(FFT)进行,FFT基函数是一种无限长的正弦函数。时频域分析适用于分析非平稳信号,其主要优点是可以在时间、频率以及幅值3个维度来观察信号的特征。栾孝驰等融合了小波包分解(WPD)、峭度值指标(KVI)与希尔伯特变换方法对滚动轴承早期故障信号进行特征识别,保留了周期性冲击成分,提高了故障信号高频分辨率,从轴承早期故障激励的高频振动信号中准确提取出了滚珠和滚棒轴承故障特征频率的1~7倍频,及其与轴转频调制的系列边频带频率等故障特征信息^[2]。

基于机器学习的故障诊断

基于机器学习的诊断方法通过使用机器学习算法,学习发动机监控数

据的特点,识别发动机的故障模式。在使用支持向量机(SVM)、决策树、K-近邻、贝叶斯等机器学习分类算法时,需要使用特征提取或特征融合技术将数据进行处理,以提高诊断结果的准确性。

支持向量机

支持向量机是基于统计学习理论的一种适合高维、小样本数据分类的学习器,是一种对数据进行二元分类的广义线性分类器。SVM可以通过核方法进行非线性分类,是常见的核学习方法之一,核支持向量机分类模型如图2所示。针对航空发动机气路故障数据样本少、非线性的问题,田静等将遗传算法(GA)参数优化的最小二乘支持向量机(LSSVM)应用于航空发动机气路故障诊断。在诊断精度和抗噪强度方面GA-LSSVM表现较好,但由于其有寻优过程,训练时间较长^[3]。

决策树

决策树属于监督学习,是一种树形结构,其中每个内部节点表示一个测量参数属性上的测试,每个分支代表一个测试输出,每个叶节点代表一种模式类别。是直观运用概率分析的一种图解法,代表的是测量参数属性与故障模式之间的一

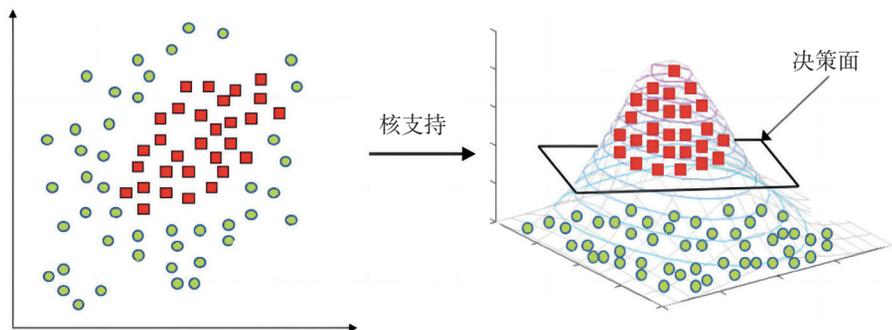


图2 核支持向量机分类模型

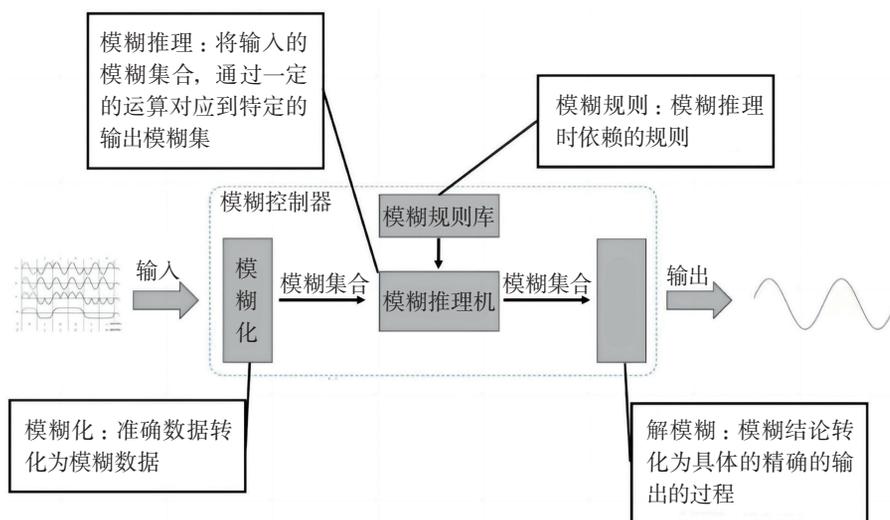


图3 模糊逻辑系统

种映射关系，对于学习到的规则能够进行可视化输出。决策树算法提取的规则可解释性强，能够减小冗余属性及噪声对决策的影响来提升诊断准确率。针对压电传感器测量的良好轴承、轴承内圈故障、轴承外圈故障以及内外圈故障4种模式振动信号，苏古马兰（Sugumaran）等将决策树与故障诊断专家知识系统相结合，将提取的统计特征形成训练特征集并输入到决策树分类器中进行训练来自动生成规则以区分轴承不同故障条件的特征，同时将规则存入知识库，后续使用知识库中的规则对故障进行分类^[4]。

模糊逻辑

模糊逻辑系统在一定条件下能以任意精度逼近给定的非线性函数。基于模糊模型的故障诊断主要有两种基本方法：一种是先建立征兆与故障类型之间的因果关系矩阵，再建立故障与征兆的模糊关系方程；另一种方法是根据先验知识建立起故障征兆与对应的原因之间的模糊规则库，利用规则库进行模糊逻辑

推理，模糊逻辑系统如图3所示。

实际应用中模糊逻辑控制系统的设计包括模糊化过程（模糊化方法、模糊集划分、模糊函数优化）、知识库（控制规则类型、一致性、完备性）、推理决策（模糊推理与解模糊方法）和优化计算（调整与优化模糊控制器）等主要内容。模糊逻辑可与神经网络进行结合，对航空发动机状态进行监视，以实现发动机健康状况的实时评估。

基于集成学习的故障诊断

机器学习的目标是学习出一个稳定的且在各个方面表现都较好的模型，但实际情况往往不这么理想，有时只能得到多个有偏好的弱监督模型。而集成学习可以组合多个弱监督模型来得到一个更好更全面的强监督模型。集成学习主要包括3类：自助聚合（Bagging）、提升法（Boosting）和堆叠法（Stacking）。

随机森林

随机森林属于Bagging方法，利用自助抽样法从发动机测量参数整

体数据集中采取有放回抽样得到 N 个数据集，在每个数据集上学习一个模型，最后的预测结果由 N 个模型投票得到，以获得更准确可靠的分类性能实现发动机故障诊断。随机森林由很多的决策树组成，每棵决策树之间没有关联。

轴间轴承作为涡轮机械的关键部件，是发生灾难性事故的主要来源。由于对冲击信号的高采样频率和高灵敏度的要求，声发射（AE）信号被广泛应用于轴间轴承故障的监测和诊断。针对轴间轴承声发射信号的非平稳和非线性问题，沈阳航空航天大学提取了多域熵作为AE信号的特征，训练建立了轴承故障诊断的随机森林模型，在轴承轴故障诊断中具有良好的泛化能力和高诊断精度^[5]。

提升法

Boosting是一种可以用来减小监督学习中偏差的算法，代表算法有：AdaBoost、梯度提升（GBDT）与XGBoost。AdaBoost算法刚开始训练时对每一个训练集赋相等的权重进行迭代训练，每次训练后，对训练失败的训练例赋以较大的权重，即利用前一轮迭代弱学习器的误差率来更新训练集的权重；与AdaBoost不同，GBDT每一次的计算是为了减少上一次的残差，算法在残差减少的方向上建立一个新的模型；XGBoost针对GBDT算法做了许多细节改进优化，包括损失函数、正则化、切分点查找算法优化、稀疏感知算法和并行化算法设计等，展现了强大的潜力。

基于深度学习的故障诊断

目前，随着传感器技术的发展，状

态监测数据的采集与获取变得更加容易。对于海量监测数据来说,传统的数理统计方法无法获取大量数据下的有用信息,而深度学习技术由于其强大的特征提取能力逐渐得到了越来越多的关注。基于深度学习的诊断方法不需要定量数学模型,可以直接利用人工智能技术,通过设备系统运行时传感器采集的状态监测数据对发动机状态进行直接建模,训练模型建立从征兆到故障源的映射过程,挖掘潜在故障关联,利用计算机学习、推理和决策实现故障诊断,缺点是判断依据没有物理含义。

核极限学习机

极限学习机(ELM)是一类基于前馈神经网络构建的方法,其特点是隐含层节点的权重为随机或人为给定的,且不需要更新,学习过程仅计算输出权重,在学习速率和泛化能力方面具有优势。ELM的一些改进版本通过引入自编码器构筑或堆叠隐含层获得了深度结构,能够进行表征学习。核极限学习机(KELM)是结合核函数所提出的改进算法,能够在保留ELM优点的基础上提高模型的预测性能。针对原始的KELM缺乏稀疏性、模型规模随样本量线性增长的问题,南京航空航天大学开发了基于递归缩减核的极限学习机(RR-KELM),简化模型结构实现KELM的稀疏性,基于可测的航空发动机参数对航空发动机故障模式具有较高识别率^[6]。

卷积神经网络

传统意义上的多层神经网络只有输入层、隐藏层、输出层,其中隐藏层的层数根据需要而定。卷积神经网络(CNN)在原来多层神经

网络的基础上加入了更加有效的特征学习部分,即在全连接层前面加入部分连接的卷积层与池化层。

目前,很多人将深度学习算法与统计分析相结合把信号识别问题转化为图像识别问题进行航空发动机智能故障诊断。西北工业大学将连续小波变换(CWT)应用于发动机控制系统传感器中的常见健康状况信号,捕获信号特征的标度图,使用经过预处理和标记的数据集训练CNN模型来提取时频图的特征,基于这些特征进行识别和隔离故障。算法具有很强的鲁棒性,准确率在97%以上^[7]。

结束语

随着科学技术的发展,传统的诊断预测维修策略已经满足不了精密性智能性发动机的要求,现代化设备对于健康管理与预测技术的要求越来越高,因此,状态检测与故障诊断技术成为了未来高性能航空发动机的关键技术之一。现有的基于数据驱动的航空发动机故障诊断研究主要分为基于数理统计的故障诊断、基于机器学习的故障诊断、基于集成学习的故障诊断和基于深度学习的故障诊断4类,在国内外均得到了广泛的研究,但是其都有特定的优缺点。基于数理统计的算法实现简单,物理含义明显,在航空发动机上应用广泛,但是算法准确率、虚警率及处理海量数据方面有待提高,基于深度学习的算法能够进行深度特征挖掘,发现特征与故障之间的潜在联系,诊断率高,但算法需要调参,依赖图形处理器(GPU)等硬件,物理含义不明确,可解释性不强,依然处于实验室验证阶段,距离工

程应用尚存在一定的差距,物理可解释性需要攻克。

航空动力

(孙泽茹,中国航空发动机研究院,工程师,主要从事发动机健康管理与控制技术研究)

参考文献

- [1] 范满意,孔祥兴,张瑞,等.一种基于统计分布特征的航空发动机气路故障融合诊断方法:CN112613186A[P].2024-07-09.
- [2] 栾孝驰,沙云东,柳贡民,等.基于WPD-KVI-Hilbert变换相结合的滚动轴承早期故障特征精准识别[J].推进技术,2022,2:43.
- [3] 田静,胡鹤翔.基于GA-LSSVM的航空发动机气路故障诊断[J].中国民航大学学报,2021,39(3):29-33.
- [4] SUGUMARAN V,RAMACHANDRAN K I. Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing[J]. Mechanical Systems & Signal Processing,2007,21(5):2237-2247.
- [5] TIAN J,LIU L,ZHANG F,et al. Multi-domain entropy-random forest method for the fusion diagnosis of inter-shaft bearing faults with acoustic emission signals[J]. Entropy,2019,22(1):57.
- [6] YOU C X,HUANG J Q,LU F. Recursive reduced kernel based extreme learning machine for aero-engine fault pattern recognition[J]. Neurocomputing,2016,6:1038-1045.
- [7] GOUL,LI H,ZHENG H,et al. Aeroengine control system sensor fault diagnosis based on CWT and CNN[J]. Mathematical Problems in Engineering,2020(6):1-12.