DOI：10.3901/JME.2024.12.090

可解释性智能监测诊断网络构造及航空发动机整机试车与中介轴承诊断应用*

王诗彬[1,2] 王世傲[1,2] 陈雪峰[1,2] 黄 海[3] 安波涛[2] 赵志斌[1,2]
刘永泉[3] 李应红[1,2]
(1. 西安交通大学航空动力系统与等离子体技术全国重点实验室 西安 710049；
2. 西安交通大学机械工程学院 西安 710049；
3. 中国航空发动机集团公司沈阳发动机研究所 沈阳 110015)

摘要：航空发动机故障预测与健康管理是提高航空发动机安全性、可靠性以及经济可承受性的关键技术。基于深度学习的人工智能方法在机械故障诊断领域受到广泛关注并开展了深入研究，但现有深度学习“黑箱算法”的现状仍然存在模型可解释性差、理论基础薄弱等问题。针对航空发动机健康管理与智能运维的迫切需求，提出航空发动机可解释性智能监测诊断网络，并在某型涡扇发动机整机长试试验中验证了异常检测与中介轴承故障诊断的有效性。将发动机振动信号先验信息融入稀疏表示模型，对模型的迭代求解算法进行展开得到结构具有可解释性的核心网络；针对航空发动机异常检测与智能诊断任务构造了基于对抗训练框架的可解释性异常检测子网络和基于特征提取框架的可解释性故障诊断子网络。本文提出的基于迭代算法展开的网络构造框架具备明确的理论基础，即网络设计有依据；稀疏表示模型驱动的可视化方法能够检验网络是否学到了与发动机故障相符的有意义的特征，即学习结果可信任。最后，通过某型涡扇发动机整机长试试验积累的超过 500 小时的试车数据，验证了本文提出的模型驱动的可解释性智能监测诊断网络在航空发动机异常检测与中介轴承故障诊断方面的有效性与可靠性。

关键词：航空发动机健康管理；可解释性人工智能；算法展开；异常检测；故障诊断
中图分类号：TH17

Interpretable Network Construction for Intelligent Monitoring and Diagnosis, and Application in Inter-shaft Bearing Diagnosis While Aero-engine Test

WANG Shibin[1,2] WANG Shiao[1,2] CHEN Xuefeng[1,2] HUANG Hai[3] AN Botao[2]
ZHAO Zhibin[1,2] LIU Yongquan[3] LI Yinghong[1,2]
(1.National Key Lab of Aerospace Power System and Plasma Technology,
Xi’an Jiaotong University, Xi’an 710049;
2. School of Mechanical Engineering, Xi’an Jiaotong University, Xi’an 710049;
3. Shenyang Engine Research Institute, Aero Engine Corporation of China, Shenyang 110015)

Abstract：Engine health management is the key technology to improve the safety, reliability and economic affordability of aero-engine. The intelligent diagnosis method based on neural networks has achieved great success in mechanical fault diagnosis, but the current network lacks the targeted design of aero-engine due to its “black box” nature, and has not been confirmed in engineering practice. In view of these problems, this paper proposes an interpretable network construction framework for intelligent diagnosis of aero-engine and verifies it in the real engine test data. The prior information of aero-engine vibration signals is integrated into the sparse representation model, and the iterative solution algorithm of the model is unrolled to obtain an interpretable core network

* 国家自然科学基金(92270111，52122504，92060302)和国家科技重大专项(J2019-I-0001-0001)资助项目。20230428 收到初稿，20240325 收到修改稿

architecture. The interpretable sub-network via adversarial training is constructed for detection tasks, and the interpretable deep feature extraction sub-network is constructed for intelligent fault diagnosis tasks. Therefore, the network architecture proposed in this paper has a clear theoretical basis, that is, ad-hoc interpretability. In addition, a visualization method is proposed to check whether the network has learned meaningful features, making it post-hoc interpretable. The characteristics of both ad-hoc and post-hoc interpretability make the network more credible when applied to aero-engine anomaly detection and fault diagnosis. Finally, in the long-term test data analysis of a real aero-engine, the interpretable network construction proposed in this paper provides an effective and credible results for fault diagnosis of inter-shaft bearings.

Key words：aero-engine prognostic and health management；algorithm unrolling；interpretable neural networks；anomaly detection；fault diagnosis

0　前言

航空发动机在军事、民用领域均有广泛的应用，自身具有很高的战略价值和经济价值，对国防和国民经济可产生巨大的效益。高推重比、大涵道比、高转速、重载是航空发动机自诞生以来主要的发展方向。这要求发动机的设计生产和材料工艺等技术同步提升，也意味着航空发动机关键部件面临巨大的可靠性考验。在军用航空发动机领域，我国进口的俄罗斯 AL-31F 发动机曾多次由于润滑系统故障导致主轴承抱死并引发空中停车，造成主力战机歼-10 频频坠毁。据不完全统计，近年来在国内服役的 AL-31F 发动机仅断轴事故就至少有 19 起，造成了严重的安全隐患与经济损失[1]。因此，提升航空发动机运行安全性成为国内外研究热点，持续保障其安全可靠运行对于提升国防实力、避免人员与财产损失具有重大意义。

主轴承作为航空发动机转子系统的支承核心，对发动机稳定运转和性能实现至关重要。在一些双转子及所有三转子发动机中，为了缩短转子长度，节省一个承力框架及相应的油腔、供回油装置等，常采用中介轴承(又称轴间轴承)的支承方式，将某一转子支承于另一转子上。与航空发动机其他主轴承不同，中介轴承的内外圈均随转子旋转，并将一个转子的载荷传递到另一个转子上。采用中介轴承支承不仅能减少零件数目并减轻发动机重量，而且可以提高整个发动机的可靠性，因此中介轴承是军用发动机和一些民用发动机中的关键基础零件。

由于中介轴承介于高压轴与低压轴之间，径向空间小，轴承的滑油供入及回油封严都比较困难，因此中介轴承常常因润滑不足而发生故障。此外，中介轴承同时连接高、低压涡轮转子，内、外环分别随内外转子旋转，因此容易引起两个转子间的动力耦合。以上是中介轴承自身结构带来的局限性，会增加其发生故障的机率。除此之外，发动机恶劣、复杂的工作环境也容易引起中介轴承故障。中介轴承通常工作在高转速、高温度和大载荷条件下，工况变化剧烈，并可能存在突然的冲击载荷和振动载荷，因此相对容易发生失效或损坏。一旦中介轴承发生故障，将直接影响航空发动机的运行安全，轻则会使转子系统振动增大，严重时甚至会导致灾难性事故。例如，中介轴承故障可能会导致发动机抱轴故障的发生，严重时会造成转子断裂带来非常严重的后果，中介轴承故障如图 1 所示。因此，开展航空发动机中介轴承的故障诊断研究具有重要意义。

图 1　中介轴承故障

在航空发动机中，中介轴承早期微弱故障有可能快速发展，导致灾难性后果，而通过定期维护提高其可靠性则需要付出高昂的成本。据国际航空运输协会统计，全球每年花费在飞机维护上的总费用可达 500 亿美元，飞机维护费用约占航空公司运营费用的 10%～15%，其中 35%～40%与发动机有关，定期更换部件是发动机维修中最重要的一项，通常可占发动机直接维修成本的 60%～70%。因此，如何在保障发动机安全运行的前提下降低维护成本，成为新一轮研究方向。航空发动机健康管理(Engine health management，EHM)是指通过机载系统和非机载系统中的传感、采集、分析、检测和数据处理等手段，提供航空发动机振动、气路、滑油、寿命等

方面的实时或近实时信息，实现状态监测、故障诊断、趋势分析和寿命管理等功能，从而预警可能影响安全运行的情况，以有针对性地安排检查维修、排除异常故障、改进功能性能、预测备件需求，进而提高航空发动机的安全性、可靠性与维修性[2]。基于航空发动机故障诊断技术形成的发动机健康管理系统是先进发动机的重要标志。

知识挖掘与智能数据分析技术获取发动机健康状态相关的状态特征，并关注未来状态的预测。尤其是以深度学习为代表的人工智能方法已经在航空发动机健康管理与智能运维中引起了广泛关注，并且在轴承等旋转机械关键零部件故障智能诊断与预测中表现了优越能力。特别是对于故障机理与故障特征明确或者故障样本与数据积累充分的机械装备。然而现有深度学习“黑箱算法”的现状仍然存在模型可解释性差、理论基础薄弱等问题，且智能诊断分析结果缺少故障机理知识支撑，无法满足航空装备高可靠与低虚警率要求。通过刷海量数据的填鸭式学习存在一系列隐患，包括运维数据样本收集的局限和偏见导致的人工智能系统的片面性，以及网络模型的“黑盒”特点，难以获得使用者的充分理解与信任。基于深度学习的人工智能方法经常从数据中学到无规律的噪声信号而难以提取与故障相关的特征，智能诊断分析结果严重缺乏故障机理相关的物理知识支撑，造成诊断结果不可靠，虚警率也无法满足航空装备健康管理系统要求。因此，如何提高航空发动机故障预测与健康管理方法的可解释性是当前迫切需要解决的问题。

在当前神经网络构造方法中，算法展开是一种被广泛研究的可解释性网络构造方法，已经被广泛研究和应用。该方法通常是通过“建模-求解-展开”的方式构建神经网络，其主要思想是通过将迭代求解算法中的迭代过程展开成一个前向计算的神经网络模型，既借助了迭代算法的计算框架，又利用了神经网络的数据驱动特性，从而实现对算法的可解释性和更高效的求解。然而，在航空发动机振动故障诊断问题中，要构建出稀疏表示模型并以此为基础构造可解释性更强的诊断网络实现深层特征提取、异常状态检测和故障诊断，仍然存在巨大的挑战。首先，航空发动机振动数据中包含一定的先验知识，可以帮助用户设计更为合理的诊断网络，然而当前的研究中较少将发动机振动信号先验与机理知识融入网络设计，以增强网络的可解释性与性能；其次，大部分关于可解释性的研究并没有构造同时具有可解释结构和可信任学习结果的网络，这使得对于航空发动机监测诊断而言，网络结构与学习结果的可信度仍然不高；最后，多数的可解释性的研究均停留在理论研究层面，尚未在实际的工程案例中对方法的有效性进行充分验证。因此，研究更为先进的、适用于航空发动机故障诊断的可解释性智能监测诊断方法迫在眉睫，对所提方法在实际的工程问题中进行验证与分析也十分具有必要性。

本文通过迭代算法展开的方式，研究航空发动机可解释性智能监测诊断网络的构造方法。将航空发动机振动信号的先验信息融入稀疏模型,推导其迭代求解公式，并通过算法展开方法构造结构上具备可解释性的核心网络，使得网络设计有依据。在核心展开网络模型的基础上，针对航空发动机异常检测和智能故障诊断任务，有针对性的设计了不同的下游网络框架。针对异常状态检测任务，构建了可解释性异常检测子网络；针对故障智能诊断任务，构建了可解释性故障诊断子网络。随后根据模型自身性质提出一套特征可视化方法，使得学习结果可信任。最后，在某型涡扇发动机整机长试试验积累的超过 500 h 的试车数据中，可解释性异常检测子网络能够指示出航空发动机寿命退化规律，并初步明确可能发生故障的时刻；可解释性故障诊断子网络具备良好的精度，通过对正常样本和异常样本的特征进行可视化分析，指示了故障部位及可能的故障类型和严重程度。

1 人工智能的可解释性

科研界对于机器学习算法“可解释性”的研究众多，但是对于其定义还未达成高度共识，很多关于可解释性的研究尚处于探索与讨论阶段。

谷歌大脑 KIM 等[3-4]认为可解释性代表了人类能够持续预测模型结果的程度，墨尔本大学 MILLER 等[5]认为可解释性是人类能够理解模型做出决策原因的程度，卡内基梅隆大学 LIPTON 等[6]认为模型的可解释性要求人类可以根据模型参数对其推理过程进行模拟。南京大学 HOU 等[7]对以上三种关于机器学习算法可解释性的理解进行了总结，并指出如果人类能够模拟模型的推理过程，或者能够持续地预测模型的输出结果，那么就可以理解模型产生决策的原因，并对模型本身和其预测结果产生信任。可见“信任”与“可解释”密切相关，香港科技大学杨强等[8]在《可解释人工智能导论》一书中对这两种概念进行了讨论，认为可解释性涉及两个方面——信任与解释。其中“信任”是人类在

社会协作中的一种心理状态，其本质上是愿意暴露自己脆弱性的表现。以航空发动机智能故障诊断为例，将人工的定期检修与排故替换为使用智能算法监测设备状态，实际上是将设备安全与飞行员生命安全交给了智能监测与诊断算法，这必须是在我们对该系统充分信任的情况下做出的决定。智能算法也需要具备与使用者解释和沟通的能力，使后者了解算法内部的决策机制从而对其产生信任，进而应用于重要场景中去。

1.1　人工智能可解释性的重要性

近些年越来越多的学者开始注意到深度神经网络等智能算法可解释性的重要性。ZHANG 等[9]提出可解释性是我们信任一个模型的基础，在测试集上的高精度不能确保模型可以正确的对特征进行编码。以色列联邦理工学院 SCETBON 等[10]认为，以往构造网络的成功经验大多基于大量的试错试验，缺少理论支撑，我们需要回归到以信号和图像处理技术为基础的网络构造方式来进一步推动计算机领域的发展。美国国防部高级研究计划局 AI 项目负责人 GUNNING 等[11]在 2019 年发表在《Science Robotics》的论文中提到，对于某些应用来说，网络的解释可能并不重要，然而对于国防、医药等领域，神经网络的可解释性至关重要。清华大学张钹院士[12]也强调，可解释、可理解是新一代人工智能技术的主要发展方向。可见神经网络的可解释性研究对于推动人工智能进一步发展、促进智能算法部署到重大装备中、保持国际科技领先地位具有重大意义。

1.2　可解释性人工智能研究现状

从理论层面出发，机器学习算法的可解释性可以分为两大类，即事前可解释性和事后可解释性[13]。这两种可解释性分别是在算法被使用之前和之后获得的，因此分别被称为事前可解释性和事后可解释性。其中事前可解释性可以通过使用本身具有可解释性的理论对特定问题进行建模来获得，由此得到的算法模型一般在结构上具有可解释性[14]。而事后可解释性则需要通过额外构建新的模型来解释，通过已有算法来实现，它通常包括大量的可视化方法，这些方法可以将算法学习到的特征进行可视化，而这些特征往往是算法分类、检测或预测的基础[15-16]。可见，算法的事前解释性有助于用户更好地理解算法是如何设计的，而事后可解释性可以帮助用户更好地了解算法如何做出决策。

事前可解释性可以通过设计将可解释性直接纳入模型结构的自解释模型来实现，其以已有知识体系驱动、与经典物理数学模型相对应、具有内在可解释性，这类算法主要包括决策树、基于规则的模型、线性模型、注意力模型等[14]。算法展开作为其中的一大分支，其发展与研究现状将在第 1.3 节专门进行详细阐述。

事后可解释性分析是可解释性人工智能研究的重点。该类研究一般将包含机器学习模型的 AI 系统看做黑盒，从而能够保证解释模型与方法和原 AI 系统之间是解耦的，保持解释层的模型无关性。事后可解释性分析的主要工作是分析模型决策依赖于指定特征的程度，因此众多学者对其进行了大量的研究工作。其中一个分支是基于梯度传播对神经网络的特征进行可视化的方法。2015 年，BACH 等[17]提出了相关性逐层传播(Layer-wise relevance propagation，LRP)，用于计算单个像素对图像分类器预测结果的贡献；2016 年，ZHOU 等[18]提出了类激活映射(Class activation mapping，CAM)，可以在任何给定图像上可视化预测的类分数，突出显示 CNN 检测到的类别相关部分；2017 年，SHRIKUMAR 等[19]提出了一种计算重要性得分的新颖方法(Deep learning important features，DeepLIFT)，其依据是根据“参考”输入的输入差异来解释某些“参考”输出的差异。另一个较大的分支为对输入单元的重要性进行归因的方法。2016 年，RIBEIRO 等[20]提出了一种局部可解释模型(Local interpretable model-agnostic explanations，LIME)，其使用可解释的局部代理模型替换决策函数，在预测周围局部学习可解释模型，以可解释和忠实的方式解释任何分类器的预测。2017 年，LUNDBERG 等[21]将 Shapley 值方法应用到特征重要性分析中，提出了一个统一的解释预测的框架(Shapley additive explanation，SHAP)，将模型预测类比为多个特征成员的合作问题，将最终预测结果类比合作中的总收益，而特征的贡献程度将决定其最终分配到的收益——重要性评估值。事后可解释性技术在不同的应用场景下适应性较强，可作为独立的可解释层构建在现有的 AI 系统之上。

1.3　基于算法展开的可解释性网络构造方法

算法展开[22]是当前被广泛研究的神经网络构造方法之一。早在 2011 年，图灵奖得主 LECUN 就发表论文，该研究将用于经典稀疏编码模型 Lasso[23]优化求解的迭代软阈值算法(Iterative soft thresholding algorithm, ISTA)在迭代方向展开为网络形态，然后使用端到端的方式自适应更新算法中的超参数，得到了一个由 ISTA 展开的可学习迭代软阈值算法(Learned iterative soft thresholding

algorithm, LISTA)，结果表明 LISTA 比经典的 ISTA 算法具有更快和更精准的特征编码能力[24]。由此可以总结出算法展开的一般步骤，首先针对先验知识进行稀疏表示建模，对模型优化求解得到迭代算法，再将迭代优化算法按照迭代方向进行展开，形成一个编码网络，然后定义网络输出与目标之间的损失函数，根据两者的残差使用网络中常用的反向传播方法自适应更新算法中原本的固定参数，如图 2 所示。通过算法展开方法所构造的网络则被称为展开网络[22]。经典的稀疏表示模型需要根据已知的研究对象先验信息进行有针对性的建模，而稀疏算法展开网络可用在研究对象机理不明确和先验信息不足的问题中，放宽对于信号结构化的约束并利用数据驱动的方式捕获信号的内在特征。

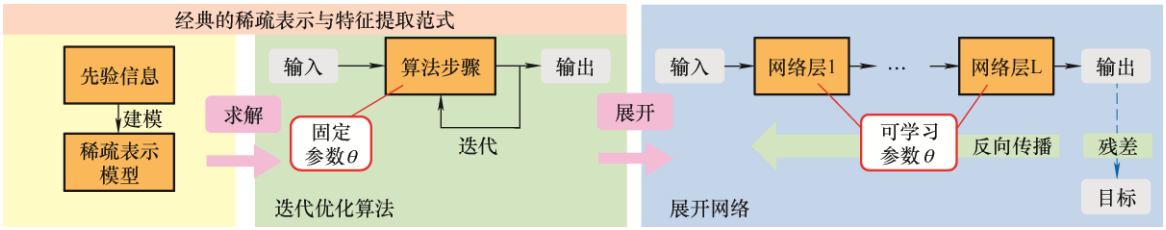


图 2 利用算法展开构造网络的过程

算法展开理论既借助了迭代算法的计算框架，又利用了神经网络的数据驱动特性，这里从函数逼近的角度对迭代算法、通用神经网络、展开网络之间的优劣势进行对比。如图 3 所示[22]，一方面，经典的迭代算法由于参数量较少，只能覆盖较小的函数子集，表示能力有限。同时，由于有理论保证，迭代算法可以在无监督的条件下逼近给定的目标函数，但是由于理论假设与实际情况不一定相符，导致迭代算法产生的结果仍存在一定偏差。另一方面，通用神经网络由于其端到端的训练方式和巨大的参数量，能够覆盖更大函数子集且能准确逼近目标函数。然而海量参数也会带来搜索空间过大、训练与推理速度过慢的问题。此外，巨大的参数量也要求使用大量样本对网络进行训练，否则会引起过拟合问题导致网络泛化性能变差。相比之下，展开网络通过扩展迭代算法的容量和使用端到端的训练方式，可以更快速、更准确地逼近目标函数。同时，展开网络在函数空间中只覆盖相对较小的子集，可以缓解训练过程中参数搜索的负担，降低训练时数据集规模的要求。总而言之，从统计学习的角度来看，迭代算法具有较高的偏差与较低的方差，而通用神经网络则有低偏差、高方差的特性，作为通用神经网络和迭代算法之间的中间形态，展开网络则可以同时拥有低偏差和低方差的优点。

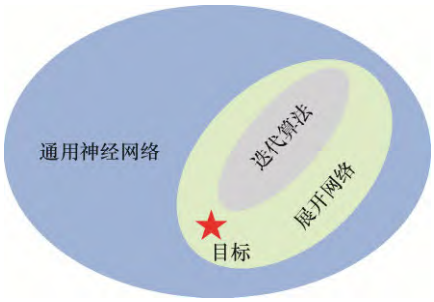


图 3 从函数逼近角度对比分析通用神经网络、迭代算法与展开网络

由于算法展开理论具有稳定的性能和良好的可解释性，近年来在多个领域获得了广泛的研究与应用。SUN 等[25]将算法展开理论应用于图像的压缩与降噪，将一个压缩感知模型的 ADMM 求解算法在迭代方向进行了展开并使用端到端的方式更新参数，试验结果表明，所提出的方法相比传统压缩感知方法，具有更快的重构速度和更好的图像重构精度；LEFKIMMIATIS[26]通过展开优化模型的梯度投影算法得到了图像降噪网络 UNLNet，该网络在参数数量远少于经典网络的条件下取得了媲美先进网络的降噪效果；DARDIKMAN-YOFFE 等[27]则在稀疏模型中构建了相关性信息，并对模型进行求解和展开，得到了 Learned SPARCOM 网络并将其应用于超分辨任务，结果显示所提出网络在处理速度上比传统深度网络快了一个数量级，并能取得更好的恢复精度与泛化能力。在理论研究层面，PAPYAN[28-29]系统性地研究了卷积神经网络与卷积稀疏表示模型之间的关系，研究表明典型的卷积神经网络中的卷积操作和非线性映射，可以对应于卷积稀疏优化算法中的矩阵相乘和软阈值映射，进一步确立了稀疏表示模型的迭代算法与神经网络之间的对应关系。MONGA 等[22]从函数逼近角度分析了算法展开理论为什么有效，即算法展开网络同时继承了迭代算法的低方差特性和神经网络的低偏差能力(端到端训练)，因此其可快速收敛到目标值。

由以上的研究发现，算法展开在应用和理论层面均得到了广泛研究，在传统的信号处理方法与神

经网络方法之间建立了宏观而又具体的联系。大部分的算法展开过程均可以归纳为“建模-求解-展开”三步，面对航空发动机中介轴承故障诊断问题，可以在建模过程中融合先验与机理知识，然后利用算法展开构造与监测对象紧密相关的诊断网络。然而当前此类研究尚未被探索，所面临的挑战与问题需要进一步分析与讨论。

2　可解释性智能监测诊断网络

本节将提出融入航空发动机振动信号先验的多层卷积稀疏编码模型，推导其迭代求解算法，将该算法进行展开得到结构具备可解释性的核心网络模型。针对航空发动机异常检测与智能诊断任务，有针对地设计不同的下游网络框架，针对异常状态检测任务构建可解释性异常检测子网络，针对故障智能诊断任务构建可解释性故障诊断子网络。最后，基于展开网络自身的特性，提出了一套事后可解释性方法。

2.1　多层卷积稀疏编码模型

航空发动机健康状态可以通过振动信号进行监测，假设测到的振动信号由可以指示设备健康状态的特征成分和无关噪声成分组成，可以表示为

$$\boldsymbol{y}=\boldsymbol{x}+\boldsymbol{e} \tag{1}$$

式中，$\boldsymbol{y}\in\mathbf{R}^N$ 是从航空发动机观察到的含噪信号，$\boldsymbol{x}$ 是关注的特征信号，$\boldsymbol{e}$ 是信号 $\boldsymbol{y}$ 中的噪声信号。假设 $\boldsymbol{x}$ 被一系列卷积字典逐层编码，即构造多层卷积稀疏编码模型，可以表示为[28]

$$\begin{aligned}\min_{\{\boldsymbol{\gamma}_i\}_{i=1}^{L}}\ \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2\quad \text{s.t.}\quad & \boldsymbol{x}=\boldsymbol{D}_1\boldsymbol{\gamma}_1, && \|\boldsymbol{\gamma}_1\|_1\leqslant t_1\\ & \boldsymbol{\gamma}_1=\boldsymbol{D}_2\boldsymbol{\gamma}_2, && \|\boldsymbol{\gamma}_2\|_1\leqslant t_2\\ & \quad\vdots && \quad\vdots\\ & \boldsymbol{\gamma}_{L-1}=\boldsymbol{D}_L\boldsymbol{\gamma}_L, && \|\boldsymbol{\gamma}_L\|_1\leqslant t_L\end{aligned} \tag{2}$$

式中，$\boldsymbol{D}_i$ 为第 i 层的卷积编码矩阵，$\boldsymbol{\gamma}_i$ 为对应的编码值，t_i 为编码值的稀疏度约束常数，$i=1,2,\cdots,L$。

根据模型(2)，多层卷积稀疏编码模型的优化目标可以表示为

$$\hat{\boldsymbol{\gamma}}_L=\arg\min_{\boldsymbol{\gamma}_L}\left\{\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{D}_{(1,L)}\boldsymbol{\gamma}_L\|_2^2+\lambda_1\|\boldsymbol{D}_{(2,L)}\boldsymbol{\gamma}_L\|_1+\cdots+\lambda_{L-1}\|\boldsymbol{D}_L\boldsymbol{\gamma}_L\|_1+\lambda_L\|\boldsymbol{\gamma}_L\|_1\right\} \tag{3}$$

式中，$\boldsymbol{D}_{(i,L)}=\boldsymbol{D}_i\boldsymbol{D}_{i+1}\cdots\boldsymbol{D}_L$，$\lambda_i$ 为平衡参数。根据式(3)可以发现 $\boldsymbol{\gamma}_L$ 可以表示编码后的高层特征。因此一旦求解出多层卷积稀疏编码模型，就可以根据编码值进行后续下游任务。

2.2　嵌套迭代软阈值算法

为了简便起见，优化目标(3)可以重写为

$$F(\gamma_L)=f\left(\boldsymbol{D}_{(2,L)}\gamma_L\right)+g_1\left(\boldsymbol{D}_{(2,L)}\gamma_L\right)+g_2\left(\boldsymbol{D}_{(3,L)}\gamma_L\right)+\cdots+g_L(\gamma_L) \tag{4}$$

式中，$f(\cdot)=\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{D}_1\cdot\|_2^2$，$g_i(\cdot)=\lambda_i\|\cdot\|_1$，$i=1,2,\cdots,L$。

近端梯度算法是用于求解稀疏编码模型的经典方法，然而优化目标(4)与经典的稀疏编码模型不同，其有 L 个正则项，传统的近端梯度算法不能直接应用与求解。因此这里首先引入梯度映射的概念用于后面的推导，梯度映射是非光滑函数的广义梯度，对于由两项组成的非光滑函数 $h(\boldsymbol{\beta})$，假设 $h(\boldsymbol{\beta})=f(\boldsymbol{\beta})+g(\boldsymbol{\beta})$，其中 $f(\boldsymbol{\beta})$ 是 Lipschitz 系数为 $1/c$ 的凸光滑函数，$g(\boldsymbol{\beta})$ 是连续但非光滑函数。则 $h(\boldsymbol{\beta})$ 的梯度映射可以表示为[30]

$$G_c^{f,g}(\boldsymbol{\beta})=\frac{1}{c}\left[\boldsymbol{\beta}-\operatorname{prox}_{cg}\left(\boldsymbol{\beta}-c\nabla f(\boldsymbol{\beta})\right)\right] \tag{5}$$

式中，$\operatorname{prox}_{cg}(\cdot)$ 是函数 $g(\cdot)$ 的近端算子[31]。由于 $g_i(\cdot)=\lambda\|\cdot\|_1$ 是一个 L_1 范数函数，那么 $\operatorname{prox}_{c_ig_i}(\cdot)=\mathcal{S}_{\lambda_ic_i}(\cdot)=\operatorname{sign}(\cdot)\max(|\cdot|-\lambda_ic_i,0)$ 是软阈值函数。令 $h_i=f+g_1+g_2+\cdots+g_{L-i}$，那么可以利用提出的广义近端梯度算法对优化方程(4)进行优化

$$\boldsymbol{\gamma}_L^{k+1}=\mathcal{S}_{\lambda_Lc_L}\left[\boldsymbol{\gamma}_L^k-c_LG_{c_L}^{h_1,g_L}\left(\boldsymbol{\gamma}_L^k\right)\right] \tag{6}$$

令 $\boldsymbol{\gamma}_{L-1}^k=\boldsymbol{D}_L\boldsymbol{\gamma}_L^k$，并根据梯度性质，优化式(6)可以重新表示为

$$\boldsymbol{\gamma}_L^{k+1}=\mathcal{S}_{\lambda_Lc_L}\left[\boldsymbol{\gamma}_L^k-c_L\boldsymbol{D}_L^{\mathrm{T}}G_{c_L}^{h_1,g_L}\left(\boldsymbol{\gamma}_{L-1}^k\right)\right] \tag{7}$$

参考式(5)可以得到 $G_{c_L}^{h_1,g_L}\left(\boldsymbol{\gamma}_{L-1}^k\right)$ 的具体形式，则式(7)可以进一步表示为

$$\boldsymbol{\gamma}_L^{k+1}=\mathcal{S}_{\lambda_Lc_L}\left(\boldsymbol{\gamma}_L^k-\frac{c_L}{c_L-1}\boldsymbol{D}_L^{\mathrm{T}}\left(\boldsymbol{D}_L\boldsymbol{\gamma}_L^k-\boldsymbol{\gamma}_{L-1}^{k+1}\right)\right) \tag{8}$$

式中，$\boldsymbol{\gamma}_{L-1}^{k+1}=\mathcal{S}_{\lambda_{L-1}c_{L-1}}\left(\boldsymbol{\gamma}_{L-1}^k-c_{L-1}G_{c_L}^{h_2,g_L}\left(\boldsymbol{\gamma}_{L-1}^k\right)\right)$。

式(8)建立了 $\boldsymbol{\gamma}_L^{k+1}$ 与 $\boldsymbol{\gamma}_{L-1}^{k+1}$ 之间的关系，从而 $\boldsymbol{\gamma}_i^{k+1}$ 与 $\boldsymbol{\gamma}_{i-1}^{k+1}$ 之间的一般关系也可以由此确定。令 $\lambda_ic_i=t_i$，$\boldsymbol{I}-c_i\boldsymbol{D}_i^{\mathrm{T}}\boldsymbol{D}_i=\boldsymbol{W}_i$，则 $\boldsymbol{\gamma}_i^{k+1}$ 与 $\boldsymbol{\gamma}_{i-1}^{k+1}$ 的嵌套优化步骤可以简化为

$$\boldsymbol{\gamma}_i^{k+1}=\mathcal{S}_{t_i}\left(\boldsymbol{W}_i\boldsymbol{\gamma}_i^k+c_i\boldsymbol{D}_i^{\mathrm{T}}\boldsymbol{\gamma}_{i-1}^{k+1}\right) \tag{9}$$

算法 1：ML-ISTA 算法

输入：

输入信号 $\boldsymbol{y}\in\mathbf{R}^N$，卷积字典 $\{\boldsymbol{D}_i\}_{i=1}^L$，正则化参数 $\{\lambda_i>0\}_{i=1}^L$，更新步长 $\{\mu_i>0\}_{i=1}^L$，迭代次数 $K\in\mathbf{N}_+$，编码层数 $L\in\mathbb{N}_+$

初始化：

$\boldsymbol{\gamma}_0^{(k)}=\boldsymbol{y},\ \forall k\in[0,K-1]$；$\boldsymbol{\gamma}_i^{(0)}=0,\ \forall i\in[0,L]$

过程：

1：主迭代 $k\in[0,K]$

2：次迭代 $i\in[1,L]$

3：$\boldsymbol{\gamma}_i^{(k+1)}\leftarrow\mathcal{S}_{t_i}\left(\boldsymbol{W}_i\boldsymbol{\gamma}_i^k+c_i\boldsymbol{D}_i^{\mathrm{T}}\boldsymbol{\gamma}_{i-1}^{k+1}\right)$

4：输出 $\{\boldsymbol{\gamma}_i\}_{i=1}^L$

因此解优化目标(3)的主要优化步骤如式(9)所示，将以上求解步骤进行总结可以得到算法 1。

此外，根据多层卷积稀疏编码模型(2)，通过下式可实现对编码值的解码过程

$$\hat{\boldsymbol{x}}=\boldsymbol{D}_1\boldsymbol{D}_2\cdots\boldsymbol{D}_L\hat{\boldsymbol{\gamma}}_L=\boldsymbol{D}_{(1,L)}\hat{\boldsymbol{\gamma}}_L \tag{10}$$

2.3 算法展开

在算法 1 中，需要预先确定一些超参数，如字典 $\boldsymbol{D}_i$、正则化参数 λ_i 和更新步长 c_i 等，其选择直接影响算法性能和信号分析效果。在稀疏编码理论中，这些超参数的选取通常需要依靠经验和先验知识进行调整和设置。然而，对于复杂或未知的数据处理问题，这些超参数的选择变得更加困难。尤其是在航空发动机中介轴承故障诊断中，难以根据故障的物理先验设置这些超参数。因此，需要进一步研究如何通过自适应方法或优化算法来自动选择超参数，以更好地应用稀疏编码理论于实际问题中。另一方面，基于固定表示字典的稀疏建模框架，无法发挥数据驱动的智能方法优势，需要根据具体问题建立针对性的模型；基于学习表示字典的稀疏建模框架，虽然可以根据数据自适应学习有效的稀疏表示字典，但是如何建立针对发动机故障本身的先验，使得稀疏表示网络能够自适应学习到针对故障的表示字典，也存在较大困难。

受到算法展开理论的启发，通过将算法 1 展开得到核心网络进行训练，从而实现了超参数的端到端训练。具体而言，将算法 1 中的核心编码公式表示为基本迭代单元，然后将该迭代单元分别在两个迭代方向(算法迭代步维度和编码层级维度)进行展开，将其转化为神经网络的形式，得到展开的核心网络结构，展开流程见图 4。此外，对解码公式(式(10))进行展开，可以得到核心网络对应的重构模块，展开流程见图 5。需要注意的是，上述两过程基于多层卷积编码模型推导而来，因而对应层的参数共享，即卷积编码字典 $\boldsymbol{D}_i^{\mathrm{T}}$ 和卷积解码字典 $\boldsymbol{D}_i$ 具有相同参数，互为转置矩阵关系，如图 6 所示。

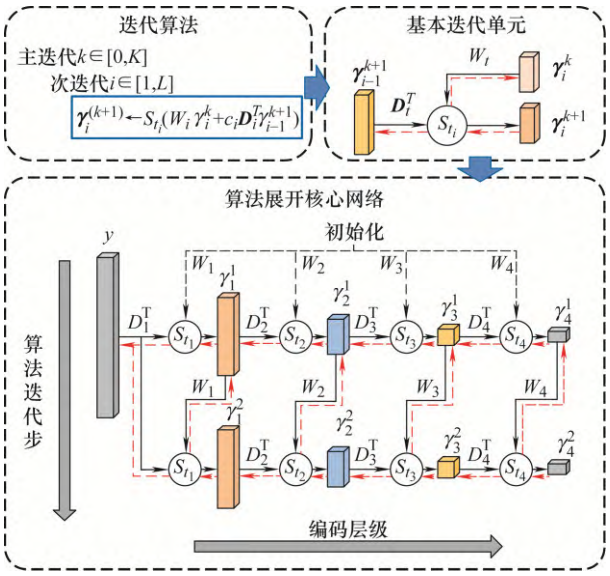


图 4 核心网络的展开流程

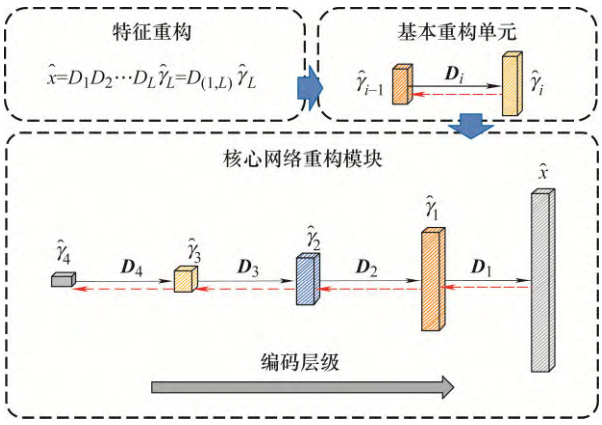


图 5 核心网络重构模块的展开流程

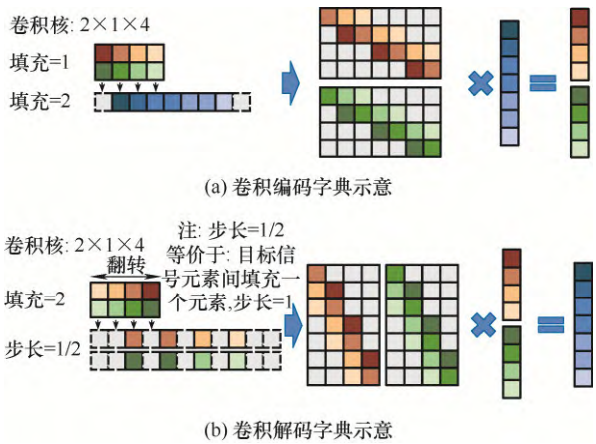


图 6 卷积操作与矩阵乘法的关系示意图

对于输入信号的编码层数，当编码层数增大，网络将会获得更好的特征提取能力，但计算复杂度显著增加，更容易出现过拟合、梯度爆炸、推理速度慢等问题，综合性能与网络训练要求，将网络的编码层数设置为 $L=4$。对于算法迭代步展开层数，当展开数增大，网络可以获得更快的收敛速度和更高的准确度，但相应的计算复杂也会变高，综合考虑数据规模与计算复杂度，本设置算法展开层数 $K=4$。网络具体参数见表 1。

表 1　算法展开核心网络结构参数

层名称	卷积核尺寸，步幅，填充	输出尺寸
卷积层 $\boldsymbol{D}_1^{\mathrm{T}}$	32×1×8，4，2	32×256
卷积层 $\boldsymbol{D}_2^{\mathrm{T}}$	64×32×8，4，2	64×64
卷积层 $\boldsymbol{D}_3^{\mathrm{T}}$	128×64×4，2，1	128×32
卷积层 $\boldsymbol{D}_4^{\mathrm{T}}$	512×128×4，2，1	512×16
转置卷积层 $\boldsymbol{D}_4$	128×512×4，1/2，2	128×32
转置卷积层 $\boldsymbol{D}_3$	64×128×4，1/2，2	64×64
转置卷积层 $\boldsymbol{D}_2$	32×64×8，1/4，5	32×256
转置卷积层 $\boldsymbol{D}_1$	1×32×8，1/4，5	1×1 024

由于核心网络是通过展开多层卷积编码模型的求解算法 ML-ISTA 所构造的，因此网络的前向传播过程相当于进行算法 1 迭代计算，也即求解多层卷积稀疏编码问题，因此网络结构上的可解释性是天然存在的。算法 1 中的超参数如字典 $\boldsymbol{D}_i$、正则化参数 λ_i 和更新步长 c_i 等作为网络参数参与到前向计算中，并通过反向传播过程进行自适应更新，具体而言，通过计算损失函数对于网络参数的梯度，并基于其梯度通过梯度下降等优化算法来更新参数。图 4 和图 5 中黑色箭头表示正向编码过程，红色箭头表示梯度的反向传播过程。

2.4　异常检测任务

为满足航空发动机监测需求，即无监督的异常检测下游任务，将算法展开核心网络应用于对抗框架和变分推理。具体而言，将核心网络及其特征重构模块分别作为异常检测子模型的编码器和解码器，组合成为一个生成器。同时，设计了一个新的鉴别器，通过对抗训练，生成器和鉴别器相互对抗，从而提高异常检测的性能。提出的可解释性异常检测子网络不仅能够有效地检测异常，而且具有较高的可解释性和稳定性。

在振动信号的异常检测问题时，理想的解决方案是根据数据分布 $P(\boldsymbol{y})$ 得到特征分布 $P(\boldsymbol{x})$，然后通过判断样本是否服从 $P(\boldsymbol{x})$ 分布来区分异常样本与正常样本。然而，当 $\boldsymbol{y}$ 是高维、强噪声干扰信号时，直接从 $P(\boldsymbol{y})$ 中学习 $P(\boldsymbol{x})$ 十分困难，因此尝试将 $P(\boldsymbol{x})$ 可以通过如下公式转换为包含低维隐藏编码 $\boldsymbol{z}$ 的形式

$$P(\boldsymbol{x})=\int_{\boldsymbol{z}}P(\boldsymbol{x}\mid\boldsymbol{z})P(\boldsymbol{z})\mathrm{d}\boldsymbol{z} \tag{11}$$

式中，$P(\boldsymbol{z})$ 为隐层编码的先验分布，$\boldsymbol{z}\in\mathbf{R}^M$ 且 $M\ll N$，$P(\boldsymbol{x}\mid\boldsymbol{z})$ 是解码概率分布。进而可以从分布 $P(\boldsymbol{z})$ 中采样，然后根据 $P(\boldsymbol{x}\mid\boldsymbol{z})$ 可以生成新的样本 $\hat{\boldsymbol{x}}$。希望隐层编码分布 $P(\boldsymbol{z})$ 可以根据数据分布 $P(\boldsymbol{y})$ 来确定，因此其可以进一步表示为

$$P(\boldsymbol{z})=\int_{\boldsymbol{y}}P(\boldsymbol{z}\mid\boldsymbol{y})P(\boldsymbol{y})\mathrm{d}\boldsymbol{y} \tag{12}$$

式中，$P(\boldsymbol{z}\mid\boldsymbol{y})$ 为编码概率分布，也称为 $P(\boldsymbol{z})$ 的后验分布。考虑式(11)和(12)，$P(\boldsymbol{x})$ 可以进一步表示为

$$P(\boldsymbol{x})=\int_{\boldsymbol{z}}P(\boldsymbol{x}\mid\boldsymbol{z})\int_{\boldsymbol{y}}P(\boldsymbol{z}\mid\boldsymbol{y})P(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z}=$$

$$\int_{\boldsymbol{z}}\int_{\boldsymbol{y}}P(\boldsymbol{x}\mid\boldsymbol{z})P(\boldsymbol{z}\mid\boldsymbol{y})P(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} \tag{13}$$

如果 $P(\boldsymbol{z}\mid\boldsymbol{y})$ 和 $P(\boldsymbol{x}\mid\boldsymbol{z})$ 是由普通网络组成的确定性函数，在这种情况下，可以根据式(13)得到一个自编码器(AE)；如果 $P(\boldsymbol{z}\mid\boldsymbol{y})$ 是一个高斯分布，并且它的均值和方差由编码器网络确定，则可以得到一个变分自编码器(VAE)。然而无论是自编码器还是变分自编码器，由于其编码和解码过程都是由普通网络实现，仍然缺乏可解释性。本文将构建一个基于对抗框架的算法展开模型来实现编码和解码过程，执行异常检测的任务。

对于 $\boldsymbol{y}\to\boldsymbol{z}$ 的编码过程，假设 $P(\boldsymbol{z}\mid\boldsymbol{y})$ 是多元高斯分布且具有对角协方差结构，则 $P(\boldsymbol{z}\mid\boldsymbol{y})$ 的表达式可以写成

$$\boldsymbol{P}(\boldsymbol{z}\mid\boldsymbol{y})=\mathcal{N}_1\left(\boldsymbol{z};\boldsymbol{\mu},\mathrm{diag}(\boldsymbol{\sigma})\right) \tag{14}$$

式中，$\mathcal{N}_1$ 表示多元高斯分布，$\boldsymbol{\mu}$ 和 $\boldsymbol{\sigma}$ 分别为分布 $\boldsymbol{P}(\boldsymbol{z}\mid\boldsymbol{y})$ 的均值和标准差，将算法展开核心网络的编码值分别输入全连接层 $\boldsymbol{D}_\mu$ 和 $\boldsymbol{D}_\sigma$ 得到这两个参数的值，从而实现将观测信号 $\boldsymbol{y}$ 映射为多元高斯分布(14)，即得到了可解释性异常检测子网络的编码过程。编码值变分映射的具体网络参数如表 2 所示。

表 2　异常检测子网络中变分映射与重采样网络层结构参数

层名称	输入尺寸	输出尺寸
全连接层 $\boldsymbol{D}_\mu$	512×16	10
全连接层 $\boldsymbol{D}_\sigma$	512×16	10
全连接层 $\boldsymbol{D}_\gamma$	10	512×16

得到 $P(\boldsymbol{z}\mid\boldsymbol{y})$ 之后，可以从 $P(\boldsymbol{z}\mid\boldsymbol{y})$ 中采样变量 $\boldsymbol{z}$，标记为 $\hat{\boldsymbol{z}}$，然后根据

$$\hat{\boldsymbol{\gamma}}_L = \boldsymbol{D}_{\boldsymbol{\gamma}}\hat{\boldsymbol{z}} \tag{15}$$

可以得到一个新采样的编码系数 $\hat{\boldsymbol{\gamma}}_L$，实现 $\boldsymbol{x}\to\boldsymbol{z}$ 的解码过程。其中 $\boldsymbol{D}_{\boldsymbol{\gamma}}\in\mathbf{R}^{N_L\times M}$ 是一个全连接层，将采样变量 $\boldsymbol{z}$ 转换到编码值相同的维度，其具体参数如表 2 所示。接着，将经过重采样的编码值 $\hat{\gamma}_L$ 输入核心网络的重构模块即可实现解码过程。因此 $P(\boldsymbol{x}|\boldsymbol{z})$ 可以由式(15)、(10)确定。如前文所述，编码过程和解码过程均由多层卷积编码模型推导而来，共享参数。

虽然在推导过程中使用了具体的多元高斯分布来表示分布 $P(\boldsymbol{z}|\boldsymbol{y})$，考虑到航空发动机振动数据分布的复杂性，希望编码器能够将数据分布 $P(\boldsymbol{y})$ 映射到任意先验分布 $P(\boldsymbol{z})$，同时保持较低的重构误差。受到对抗自编码器网络结构的启发，这里使用对抗训练策略来执行变分推理，将构造的编码器和解码器组合成一个生成器，并为对抗训练设计了一个新的鉴别器。鉴别器的存在可以实现将数据分布映射到任意的先验分布，也有助于异常检测过程中实现自动异常样本判别，而无需手动设置异常判断标准。此时得到了一个可解释性异常检测子网络，如图 7 所示。这里判别器通过三个串联的全连接层(FC_1，FC_2，FC_3)实现，具体结构参数设置如表 3 所示。

表 3 异常检测子网络中鉴别器网络结构参数

层名称	输入尺寸	输出尺寸
全连接层 FC_1	10	512
全连接层 FC_2	512	256
全连接层 FC_3	256	1

对抗框架下的训练可以分为两个阶段，第一阶段是生成器的优化，在这个阶段定义了对抗损失与重构损失两个损失。对抗损失函数定义为

$$\mathcal{L}_{adv} = \mathbf{E}_{y\sim p_y}\left\{-\ln\left[f\left(G_e(\boldsymbol{y})\right)\right]\right\} \tag{16}$$

式中，$\mathbf{E}_{y\sim p_y}\left\{-\ln\left[f\left(G_e(\boldsymbol{y})\right)\right]\right\}$ 表示判别器结果 $-\ln\left[f\left(G_e(\boldsymbol{y})\right)\right]$ 关于分布 $\boldsymbol{p}_{\boldsymbol{y}}$ 的数学期望，$f(\cdot)$ 是判别器函数，G_e 是生成器中的编码器，$p_y=\boldsymbol{P}(\boldsymbol{y})$。重构损失函数为

$$\mathcal{L}_{con} = \mathbf{E}_{y\sim p_y}\|\boldsymbol{y}-G(\boldsymbol{y})\|_1 \tag{17}$$

式中，G 表示生成器函数，则生成器的综合损失函数为

$$\mathcal{L}_G = w_{adv}\mathcal{L}_{adv} + w_{con}\mathcal{L}_{con} \tag{18}$$

式中，w_{adv} 和 w_{con} 是调整对抗损失和重构损失比重的权重参数。第二阶段是判别器的优化，其损失函数为

$$\mathcal{L}_D = \mathbf{E}_{y\sim p_y,\boldsymbol{z}\sim p_{\boldsymbol{z}}}-\left\{\ln\left[1-f\left(G_e(\boldsymbol{y})\right)\right]+\log\left[f(\boldsymbol{z})\right]\right\} \tag{19}$$

式(18)和(19)中分别得到了生成器和判别器损失函数，然后就可以通过交替优化这两个损失函数来实现整个网络的参数更新。

2.5 智能故障诊断任务

为了应对有监督的故障诊断下游任务，将算法展开核心网络应用于特征提取框架。核心网络作为特征提取器，并构造特征分类器。在分类器中首先在特征提取器输出编码 γ_L^K 的基础上添加全局最大池化(Global max-pooling，GMP)层实现对编码组稀疏的间接约束，同时也可以进一步滤除编码中的噪声信息。然后在此基础上添加了一个全连接(Fully connected，FC)层和一个 Softmax 层用于故障分类，分类器网络具体结构参数见表 4。最终得到了一个可解释性故障诊断子网络，如图 7 所示。

表 4 特征分类器网络结构参数

层名称	输入尺寸	输出尺寸
全局最大池化层 GMP	512×16	512
全连接层 FC	512	故障类别数

2.6 事后可解释方法

为了验证多层卷积稀疏编码展开网络能够从数据中学到有物理含义的特征，本文提出一套针对该网络的多尺度特征可视化方法，用于事后可解释性分析，使得航空发动机异常检测与故障诊断的结果更加可信任。该方法包括两种可视化原理，即整体特征可视化和原子特征可视化。

首先是整体特征的可视化分析方法。基于式(10)以及核心网络重构模块输出的特征信号 $\hat{\boldsymbol{x}}$，也就是异常检测子网络解码器输出特征和智能诊断子网络的重构特征，如图 7 所示，通过对 $\hat{\boldsymbol{x}}$ 进行可视化并进行分析来验证学习结果的可解释性。在仿真案例中，$\hat{\boldsymbol{x}}$ 可以与真实特征(ground-truth)进行比较，以判断展开网络是否可以从含噪信号中恢复出有意义的特征。在工程案例中，可以分析 $\hat{\boldsymbol{x}}$ 是否与航空发动机主轴承故障特性相匹配来判断特征是否有意义，使得分析结果值得信任。

其次是原子特征可视化分析方法。考虑式(2)、(3)，特征的估计值 $\hat{\boldsymbol{x}}$ 可以重写为

$$\hat{\boldsymbol{x}} = \boldsymbol{D}_1\boldsymbol{\gamma}_1 = \boldsymbol{D}_{(1,2)}\boldsymbol{\gamma}_2 = \cdots = \boldsymbol{D}_{(1,L)}\boldsymbol{\gamma}_L \tag{20}$$

已知 $\hat{\boldsymbol{x}}$ 是从含噪信号 $\mathbf{y}$ 中恢复的特征信号，其

可以被一系列字典 $\boldsymbol{D}_1,\boldsymbol{D}_{(1,2)},\cdots,\boldsymbol{D}_{(1,L)}$ 进行表示，因此展开网络前 j 层学习到的原子特征反映在字典 $\boldsymbol{D}_{(1,j)}$ 的原子中，该字典可以通过下式计算得到

$$\boldsymbol{D}_{(1,j)}=\boldsymbol{D}_1\boldsymbol{D}_2\cdots\boldsymbol{D}_j \tag{21}$$

式中，$j=1,2,\cdots,L$。单个卷积字典 $\boldsymbol{D}_i$ 可以根据网络结构参数如卷积核尺寸、步幅、填充等来确定。在确定单个卷积字典 $\boldsymbol{D}_i$ 之后则可以根据式(21)得到综合卷积字典 $\boldsymbol{D}_{(1,j)}$，从而可视化网络学到的前 j 层的原子特征。

因此，提出了两个可视化原则，两种特征可视化方法是从多个尺度进行的，能够帮助用户更好地理解和信任检测结果，使其更加可信。

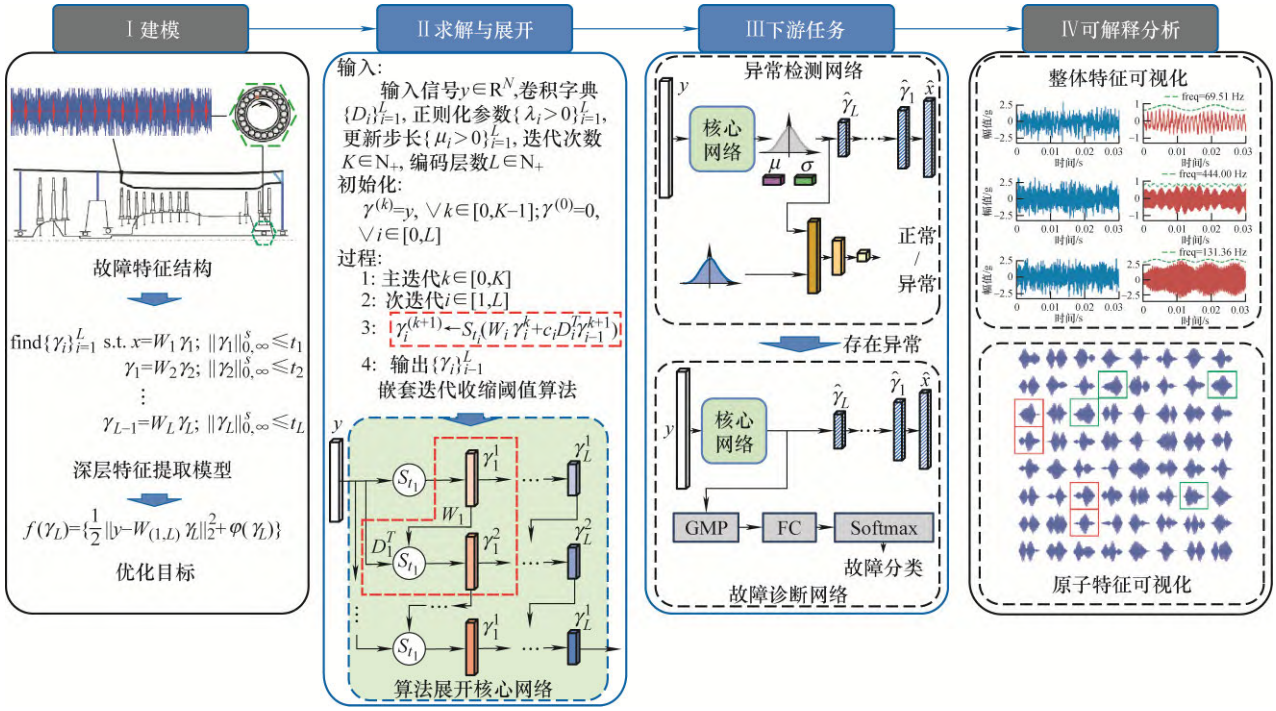


图 7　可解释性智能监测诊断网络构造流程示意图

3　航空发动机整机试验分析

为了研究航空发动机长期运行过程中的部件失效规律，开展了某型航空发动机的台架状态长试试验，其中发动机结构示意图如图 8 所示。在试车过程中，分别在 3 号支点所在截面的机匣安装边附近安装了振动传感器，采样频率为 32 768 Hz，用于监测转子-轴承系统的振动情况，其中 V3 表示支点 3 所在截面附近的振动测点，发动机测点与通道信息

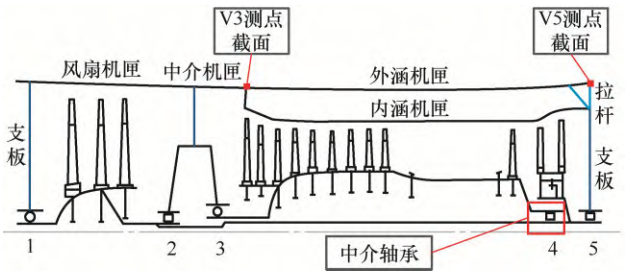


图 8　发动机结构简图

见表 5。试车试验总历时三个半月，实际有效试车总时长约为 500 h。

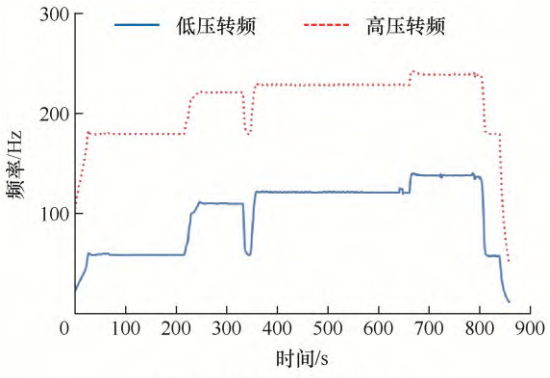
试车过程中进行基于滑油系统的在线监控，即以滑油系统中铁屑的含量作为监控指标，监测系统在本次试车过程中并未报警。由于整机试车条件限制，无法对轴承损伤情况进行在线监测，试验后分解检查发现，4 号轴承(中介轴承)存在明显的磨损情况，全部滚子边缘有环带状剥落，内圈滚道整体接触痕迹较重，见图 9。试验后复查滑油系统监控参数，试车过程中，光谱检查结果无异常，滑油滤、磁塞上未发现轴承材料金属屑。可见基于滑油系统的监测方法存在一定局限，开展基于振动分析的监测方法具有一定的工程意义。针对该情况运用可解释智能监测诊断网络进行中介轴承振动信号分析。从 500 h 的整机试验数据中综合考虑前期、中期、后期不同阶段，选取了具有代表性的 23 组试车数据进行进一步分析，尝试利用提出的可解释性异常检测与智能故障诊断方法对发动机中介轴承性能退化规律和故障损伤程度进行深入研究。

图 9 发动机中介主轴承滚子和内外圈损伤情况

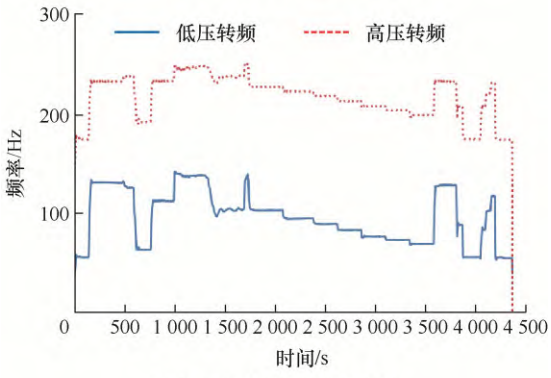
表 5 发动机测点与通道信息

通道编号	数据类型	测点代号	传感器型号	测点位置	主轴承支点
CH3	振动	V3	7537	中介机匣后安装边	3 支点
CH9	转速	N1	—	低压转子转速	—
CH10	转速	N2	—	高压转子转速	—

根据发动机试车的要求，各次试车的试车程序并不完全一致，甚至存在较大区别。例如图 10 给出了具有代表性的两组试车过程中的高低压转频变化情况。可以看出，试验过程中每次的试车程序和试车时数均存在明显差异，既包括图 10a 的阶梯式逐步上推和快速下拉程序，也包括图 10b 中的快速上推和缓慢阶梯式下拉的试车程序。针对这种情况，研究多次试车过程中的部件性能的退化规律，则需要在不同试车试验中寻找统一的基准工况，并在同一基准工况下在开展数据演化规律分析才能得到较为可靠的结论。



(a) 第1次试车高低压转频变化



(b) 第14试车高低压转频变化

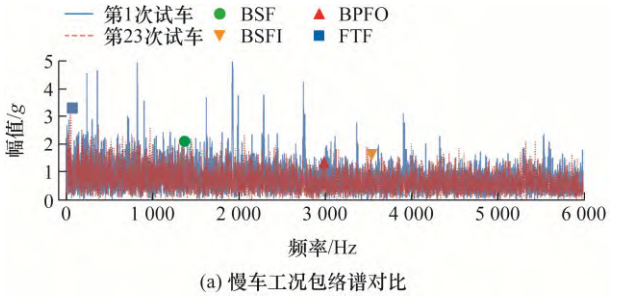
图 10 试车过程中的高低压转频变化

3.1 数据清洗与初步分析

首先统计了选取的 23 次试车过程中归一化高压转速的综合数据分布，发现高压转频分别在航空发动机的慢车工况、暖车工况和巡航工况下有密集分布，与发动机实际工作状态相符。因此数据分布主要以这 3 种工况为基准，分析试验过程中发动机中介轴承的性能退化规律。这里使用包络谱分析观察在试车过程中的频率成分变化规律。每种工况下中介轴承各部件的故障特征频率如表 6 所示。选取在 3 种工况下第 1 次试车和第 23 次试车振动信号的包络谱进行对比，结果如图 11 所示(其中 BSF、BPFO、BPFI 和 FTF 分别表示滚子旋转频率、外圈滚子通过频率、内圈滚子通过频率和保持架旋转频率)。可以发现，在慢车和暖车工况下，第 1 次试车和第 23 次试车包络谱变化较小，尤其是在不同故障位置对应的故障频带附近，两次试车之间并没有明显差异性，难以判断故障类型。在巡航工况下，内圈故障特征频率附近出现了差异化，第 23 次试车时该频带附近频率值明显高于第 1 次试车，但是由于该频率没有完全与轴承内圈故障特征频率对应，并且外圈故障特征频率在其他两种工况下并没有明显发生变化，因此仅靠包络谱分析同样难以量化对比试车过程中中介轴承性能退化规律或确定可能发生的故障类型。

表 6 中介轴承各部件故障特征频率

工况	滚动体/Hz	内圈/Hz	外圈/Hz	保持架/Hz
慢车	1 392.24	3 598.39	3 037.61	69.51
暖车	1 956.18	5 055.97	4 268.06	69.57
巡航	2 249.90	5 815.13	4 908.87	65.68



(a) 慢车工况包络谱对比

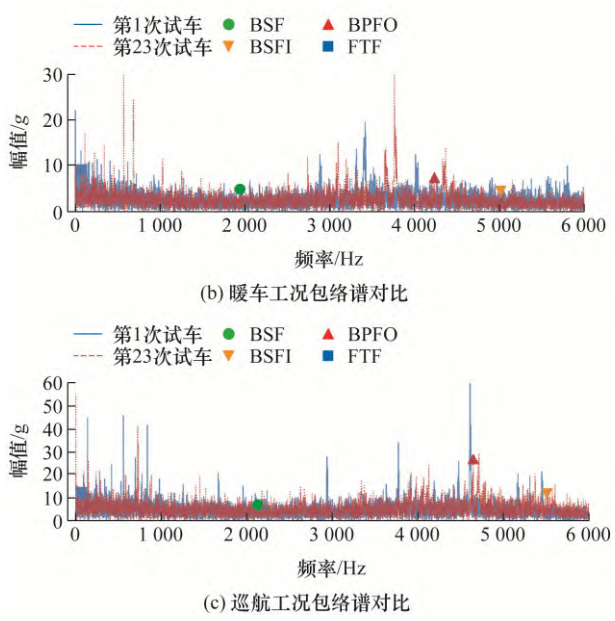


图 11　在 3 种工况下第 1 次试车和第 23 次试车振动信号包络谱对比

3.2　中介轴承性能退化分析与异常检测

3.2.1　异常趋势分析

本节借助异常检测方法分析发动机试车过程中中介轴承的性能退化规律，尝试分析可能发生故障的时间点。因为无法确认具体异常发生的时间点，假设前 5 次试车的数据为正常样本(由于是 500 h 的整机长试，且初始状态是完好的发动机，因此前 5 次大概率是正常发动机)，以慢车工况下的振动数据训练异常检测模型，然后将从第 6 次试车开始的每一次试车数据全部标记为异常样本(这里也是假设异常，具体是否异常，可以从分析结果中观察)，测试可解释监测诊断网络中的异常检测子网络，并选择变分自编码器(Variational autoencoder, VAE)，GANomaly[33] 和对抗自编码器(Adversarial autoencoders, AAE)作为对比方法。四种方法的检出率变化如图 12 所示，可以发现，随着试车次数的增加，异常检测子网络对于样本的异常检出率逐渐增大，从第 10 次试车开始检出率基本稳定在 100%附近。可见异常检测子网络对于发动机试车过程中的部件性能退化过程较为敏感，具有符合性能退化规律的检出率变化趋势。此外，在工程问题中，多次检测虚警率小于 5%则可以认定检测结果具有稳定性，因此可以初步认定在第 10 次试车时发动机状态产生了较为明显的异常，暂时不能判断出具体的异常位置或故障类型。而对比方法 VAE，AAE 和 GANomaly 的检出率变化较为杂乱，难以捕获中介轴承微弱的故障特征，提取的特征中混杂了大量的无关信息，导致没有明显的规律，从而对比证明了所提方法对于发动机异常状态的敏感性和异常检测的有效性。需要注意的是，在试验过程中由于故障的发生往往存在一定的渐变演化过程，因此在本节和后续章节中提到“正常”、“异常”以及“故障”均为相对概念。

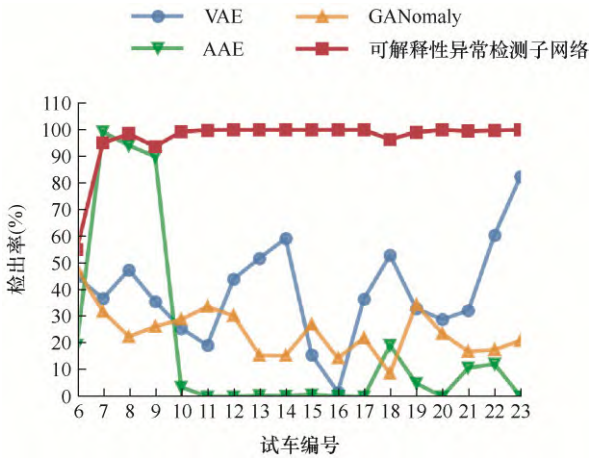


图 12　可解释性异常检测子网络及对比方法检出率变化规律

表 7　在 3 种工况下异常检测性能对比

工况	方法	检出率(TPR)	虚警率(FPR)	准确率(ACC)	AUC	F1 分数
慢车	VAE	97.44%	0.34%	98.55%	0.998 6	0.985 3
	GANomaly	100.00%	32.80%	83.60%	0.791 7	0.859 1
	AAE	20.34%	6.00%	57.17%	0.660 8	0.321 9
	可解释性异常检测子网络	99.88%	0.08%	99.90%	0.999 8	0.999 0
暖车	VAE	72.20%	24.20%	73.58%	0.828 4	0.770 8
	GANomaly	92.70%	12.80 %	90.58%	0.919 6	0.923 7
	AAE	0.05%	19.32%	31.06%	0.054 3	0.000 8
	可解释性异常检测子网络	93.12%	2.04%	94.98%	0.985 6	0.958 0
巡航	VAE	81.38%	9.81%	85.57%	0.913 2	0.855 2
	GANomaly	96.45%	10.60%	90.82 %	0.896 5	0.809 7
	AAE	13.92%	1.26%	81.57 %	0.319 9	0.234 2
	可解释性异常检测子网络	96.45%	4.66%	95.55%	0.942 0	0.897 8

3.2.2 异常检测性能对比分析

在第 3.2.1 节中验证了模型的异常趋势分析能力，其在异常状态振动的实时性上表现良好。为了定量分析可解释性异常检测子网络对于发动机异常状态的检测性能，进行异常检测性能的对比。由于中间试车过程中，轴承状态未知，仅通过最后一次试车后的分解得知中介轴承发生滚子与内圈损伤，而大量中间样本状态未知。为严谨起见，这里选择前五次试车数据作为正常样本(标签为 0)，第 23 次试车数据为异常样本(标签为 1)，其中正常样本按照 4∶1 划分训练集和测试集，异常样本则全部划入测试集。本节所对比的方法同样包括 VAE、GANomaly 和 AAE。这里将不同方法对于样本的异常概率预测值分布进行统计,计算四种方法的检出率(True positive rate, TPR)、虚警率(False positive rate, FPR)、准确率(Accuracy, ACC)、ROC 曲线下面积(Area Under Curve, AUC)和 F1 分数[32]，结果如表 7 所示。从表中可以看出，在检出率方面，GANomaly 与可解释性异常检测子网络性能接近，均优于其他方法，而在虚警率方面，可解释性异常检测子网络明显优于 GANomaly。在航空发动机监测诊断领域，虚警率是非常重要的指标，其重要程度甚至高于检出率，因为虚警所导致的额外检修带来的经济损失往往是巨大的。准确率、F1 分数和 AUC 均综合考虑了方法的检出率与虚警率，在这三项综合指标上，可解释性异常检测子网络明显优于其他方法。

因此，总体而言提出的可解释性异常检测子网络与现有方法相比，在异常检测性能上具有更优秀的表现。这与其网络结构的可解释性有关，网络结构具备可解释性使网络在针对具体的异常检测任务时可以获得更好的检测性能。

3.2.3 特征可视化分析

这里对异常检测子网络从正常样本中学到的特征进行可视化，来说明检测结果是否可信任。

首先是总体特征可视化。这里选取完成训练的可解释性异常检测子网络，然后将测试数据中的正常样本输入模型中，观察该网络从数据中提取的特征，并对其进行可视化，结果如图 13 所示。这里选取了三种工况下的样本重构特征，可以发现，在慢车和巡航工况下，可解释性监测诊断网络中的异常检测子网络可以从数据中重构出中介轴承保持架转频和其二倍频，在暖车工况下则重构出了高压转频的二倍频。考虑到可视化过程中输入样本均为正常样本，且该网络在训练时也仅有正常样本用于模型参数更新，可以说明中介轴承保持架转频和高压转子转频均为发动机在“正常”运转时就会出现的频率。其高低压转频作为振动信号中的主要频率成分之一，无论发动机工作状态正常或异常均会出现，因此在重构特征中出现高压转子转频表明网络从数据中学到了设备正常运转时的主要频率成分。而在发动机“正常”工作时，网络重构出了中介轴承保持架转频，其原因可能是中介轴承相比其他主轴承更容易润滑不充分[34]。

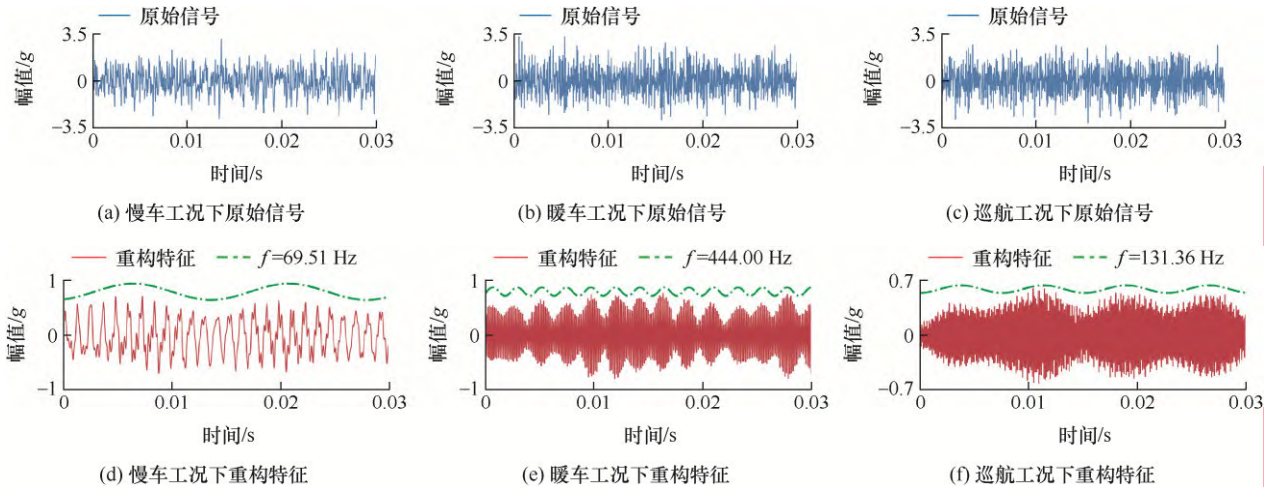


图 13 可解释性异常检测子网络重构特征可视化

接着进行原子特征的可视化分析。以暖车工况为例，对可解释性异常检测子网络的字典原子进行了观察和分析。由于相关的特征频率较低难以在长度较短的字典上表现出来，只关注长度最长的 $\boldsymbol{D}_{(1,4)}$ 字典原子并进行观察和分析，部分原子如图 14 所示(红色对应高压转频，绿色对应低压转频)。由于字典原子长度较小，而采样频率较高，因此无法采用傅里叶变换等方法进行精细化频域分析，本研究基

于字典原子的大致周期来区分不同类型转频。通过观察这些原子，可以发现网络学到了前期试车信号中高低压转频相关的特征。基于这些特征，该网络能够对发动机异常状态进行检测，同时提高了其学习结果的可信度。

上述可视化方法，验证了提出的可解释性异常检测子网络在具备良好检测性能的同时，也具备了良好的可解释性，即其学习结果是可信的。

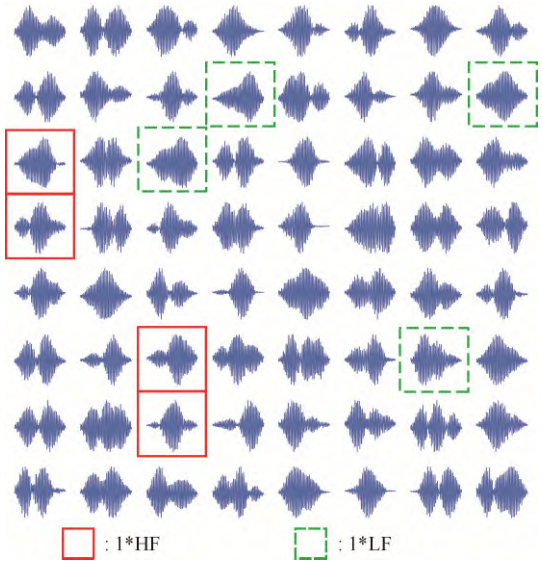


图 14　可解释性异常检测子网络中 $\boldsymbol{D}_{(1,4)}$ 字典的部分原子

3.3　中介轴承故障智能诊断

3.3.1　故障诊断性能对比

本节使用可解释性监测诊断网络中的故障诊断子网络进一步开展故障诊断试验研究。与上节研究同理，考虑到中间试车轴承状态未知，这里使用第 1 次试车和第 22 次试车在三种工况下的振动数据作为训练集对网络进行二分类训练，分别标记第 1 次试车数据为正常(标签为 0)，第 22 次试车数据为故障(标签为 1)。将第 2 次试车和第 23 次试车数据分别标记为正常和故障，并且以此作为测试集对比包括可解释性故障诊断子网络、循环神经网络(Recurrent neural network, RNN)、长短期记忆递归神经网络(Long short term memory, LSTM)、残差神经网络(Resdual network, ResNet)和 SincNet [35]在内的故障诊断方法性能。几种方法的故障诊断诊断准确率如表 8 所示。由于试车初始阶段与结尾阶段数据的区分性本身就较大，二分类任务难度低，以上方法均能达到接近 100%的精度。尤其是 ResNet-18 方法对于数据的拟合能力最强，很容易获得 100%的诊断精度。

表 8　在 3 种工况下异常检测性能对比

工况	RNN	LSTM	SincNet	ResNet-18	可解释故障诊断子网络
慢车	0.998 9	0.996 8	0.989 3	1.000 0	0.999 9
暖车	0.999 5	0.999 0	0.981 4	1.000 0	0.999 6
巡航	0.995 9	0.995 2	0.772 3	0.999 4	0.998 1

3.3.2　特征可视化分析

在第 3.3.1 节中对网络的性能进行验证，由于试车条件限制，试验数据集选择状态明确的样本，这也导致模型在表现上过于乐观。但借助本文网络天然的编解码结构以及特征提取能力，可以对故障特征进行进一步分析，从而实现中介轴承故障隔离。本节应用整体特征和原子特征的可视化方法对可解释性故障诊断子网络进行分析，检测网络学到的特征是否可以指示故障，以及是否可以指示轴承退化规律。

首先是整体特征可视化分析，选取测试数据输入训练完成的可解释性故障诊断子网络中，从输入数据中提取的特征进行重构，从而判断诊断结果的可靠性和中介轴承可能的故障类型。为了消除个别样本包络分析所带来的随机性，进一步求取了所有正常样本和异常样本重构特征的平均包络谱图，结果如图 15 所示。可以看出，在平均谱图下慢车工况

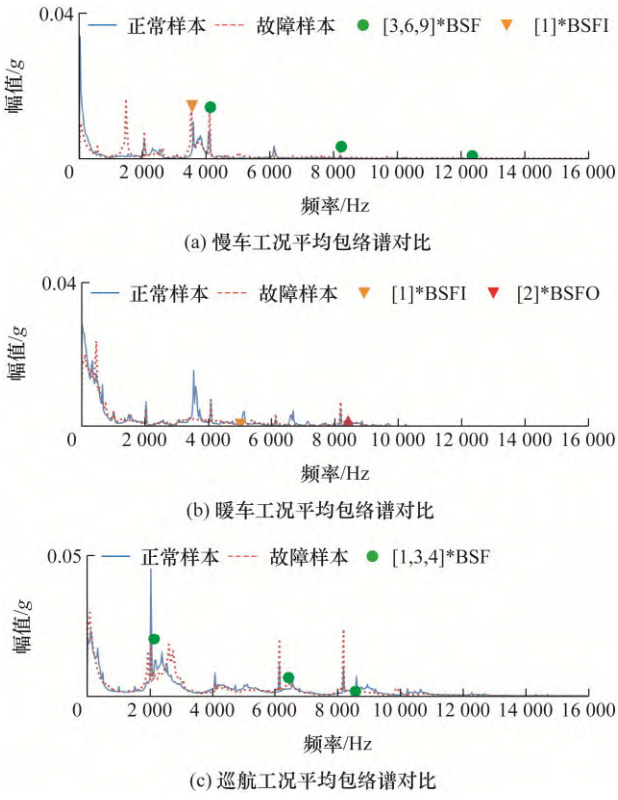


图 15　在 3 种工况下第 1 次试车和第 23 次试车重构特征平均包络谱对比

和巡航工况故障样本相比正常样本均产生了较为明显的滚子故障特征频率 BSF 的倍频，暖车工况下则出现了内圈故障特征频率 BSFI 和外圈故障特征频率 BSFO 的倍频，由此得出的结论与图 13 一致，并且由于使用了平均谱图，结论更具有确定性。

中介轴承所有滚动体均发生了严重的磨损故障，轴承内外圈滚道也有不同程度的损伤。通过对比图 11 和图 15 可以发现，在原始信号的包络谱中并不能发现明显的故障特征频率，而在经过整体特征可视化方法重构出特征之后，包络谱中出现了较为明显的故障特征频率及其倍频。

为了研究滚动体故障特征随试车次数变化的规律，这里计算每次试车数据原始信号以及重构特征的平均包络谱中 3、6、9 倍滚动体故障特征频率成分的能量占整个包络谱能量的比例，作为故障特征的能量占比，并绘制了试车次数与滚子故障特征频率能量占比之间的关系，结果如图 16 所示。可以看出，随着试车次数的增加，重构特征中滚子故障特征频率的 3、6、9 倍频率能量占比也会波动上升。特别是在第 10 次试车之后，出现了明显上升，这与发动机异常检测结果相符。而原始信号中滚子故障特征频率的能量占比随时间变化不明显，说明可解释性故障诊断子网络具有出色的特征提取能力。

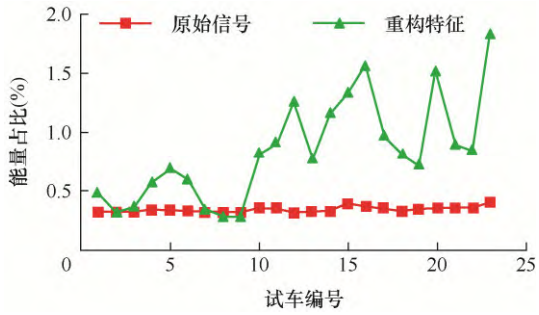


图 16 慢车工况下滚子故障特征频率能量占比随试车的变化

接下来进行字典原子特征的可视化分析，以慢车工况下训练的网络为例，对其中从 $\boldsymbol{D}_{(1,2)}$ 到 $\boldsymbol{D}_{(1,4)}$ 字典的部分原子进行可视化分析，具体见图 17。由于字典原子长度和采样频率限制，导致频谱分析分辨率过大，这里主要是基于重构特征中相对明显的频率特征对字典原子进行大致筛选。可以看出，可解释性故障诊断子网络可以学习不同尺度的成分特征，主要的故障特征(如滚子故障特征频率)和主要的正常特征(如高压转频)，都被很好地学习，说明了该网络具备良好的可解释性，其学习结果值得信任。除了框起来的原子特征外，其他的原子特征仍可能具有一定的物理信息，但由于噪声的混叠等因素，表现相对不明显，但仍是构成重构特征的成分之一。

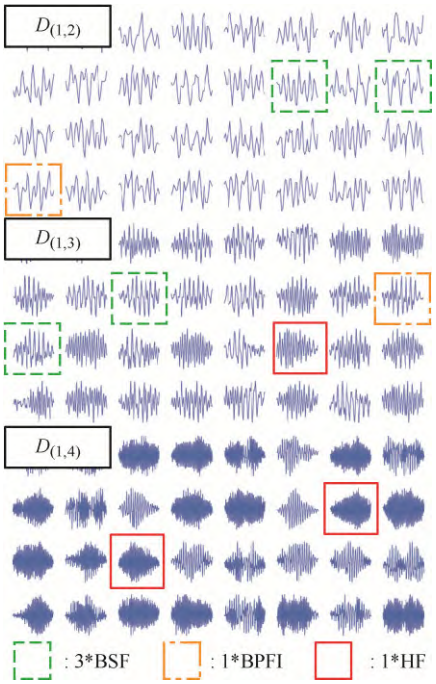


图 17 可解释性故障诊断子网络中从 $\boldsymbol{D}_{(1,2)}$ 到 $\boldsymbol{D}_{(1,4)}$ 字典的部分原子

通过性能验证和可视化分析，充分验证了可解释性故障诊断子网络在故障诊断性能和可解释性方面相比于其他方法的优势，证明了该方法在实际航空发动机关键部件智能故障诊断中的有效性。

4 结论

针对航空发动机健康管理与智能运维的迫切需求，本文提出了航空发动机可解释性智能监测诊断网络，为实现更高效、可解释和可定制化的神经网络提供了一种新的思路和方法，并在某型涡扇发动机整机长试试验中验证了异常检测与中介轴承故障诊断的有效性，并得出以下结论。

(1) 基于“建模-求解-展开”的算法展开理论，针对航空发动机振动数据先验构建了多层卷积稀疏编码模型，对模型进行求解与展开，得到了结构具备可解释性的核心网络，而从使得满足航空发动机监测诊断需求的网络设计有依据。针对航空发动机中介轴承的异常检测和故障诊断任务，将核心网络应用于对抗网络框架和特征提取框架，分别构建了可解释性异常检测子网络和故障诊断子网络，并针对核心网络提出了特征可视化方法，从而保证学习结果可信任。

(2) 所提出的异常检测与故障诊断子网络相辅相成，用于某型航空发动机整机长试验数据分析，异常检测子网络可以指示航空发动机出现异常的时刻，且异常检测性能优于其他对比方法；通过特征可视化分析初步明确发动机早期试车中的信号构成。随后通过智能故障诊断子网络对正常样本和异常样本的重构特征和原子特征进行分析，明确了中介轴承可能的故障类型和严重程度，验证了结论的正确性，更说明了本方法能够为航空发动机健康管理与智能运维提供可信依据。

参 考 文 献

[1] 向巧，张铀，许亚平. 三型涡扇发动机故障模式与机理分析及预防技术[M]. 北京：航空工业出版社，2016.
XIANG Qiao，ZHANG Zhou，XU Yaping. Failure mode and mechanism analysis and prevention technology in Type III turbofan engine[M]. Beijing：Aviation Industry Press，2016.

[2] 陈雪峰，王诗彬，曹明. 航空发动机快变信号分析及故障诊断系统[M]. 北京：科学出版社，2022.
CHEN Xuefeng，WANG Shibin，CAO Ming. Aero-engine rapid change signal analysis and fault diagnosis system[M]. Beijing：Science Press，2022.

[3] KIM B，KHANNA R，KOYEJO O O. Examples are not enough，learn to criticize! criticism for interpretability[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems，2016：2288-2296.

[4] DOSHI-VELEZ F，KIM B. Towards a rigorous science of interpretable machine learning[J/OL]. [2024-03-12]. https://arxiv.org/abs/1702.08608.

[5] MILLER T. Explanation in artificial intelligence：Insights from the social sciences[J]. Artificial Intelligence，2019，267：1-38.

[6] LIPTON Z C. The mythos of model interpretability：In machine learning，the concept of interpretability is both important and slippery[J]. Queue，2018，16(3)：31-57.

[7] HOU B J，ZHOU Z H. Learning with interpretable structure from gated RNN[J]. IEEE Transactions on Neural Networks and Learning Systems，2020，31(7)：2267-2279.

[8] 杨强，范力欣，朱军. 可解释人工智能导论[M]. 北京：电子工业出版社，2022.
YANG Qiang，FAN Linxin，ZHU Jun. Introduction to explain artificial intelligence[M]. Beijing：Electronic Industry Press，2022.

[9] ZHANG Q，WU Y N，ZHU S C. Interpretable convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition，2018：8827-8836.

[10] SCETBON M，ELAD M，MILANFAR P. Deep k-svd denoising[J]. IEEE Transactions on Image Processing，2021，30：5944-5955.

[11] GUNNING D，STEFIK M，CHOI J，et al. XAI—Explainable artificial intelligence[J]. Science Robotics，2019，4(37)：7120.

[12] 张钹. 展望人工智能的未来[J]. 科学世界，2020(4)：85.
ZHANG Bo. Prospect the future of artificial intelligence[J]. Scientific World，2020(4)：85.

[13] DU M，LIU N，HU X. Techniques for interpretable machine learning[J]. Communications of the ACM，2019，63(1)：68-77.

[14] CHEN X，DUAN Y，HOUTHOOFT R，et al. Infogan：Interpretable representation learning by information maximizing generative adversarial nets[C]// Advances in Neural Information Processing Systems，2016：29.

[15] ZEILER M D，FERGUS R. Visualizing and understanding convolutional networks[C]//Computer Vision–ECCV 2014：13th European Conference，Zurich，Switzerland，September 6-12，2014，Proceedings，Part I 13. Springer International Publishing，2014：818-833.

[16] ZHANG Q，ZHU S C. Visual interpretability for deep learning：A survey[J]. Frontiers of Information Technology & Electronic Engineering，2018，19(1)：27-39.

[17] BACH S，BINDER A，MONTAVON G，et al. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation[J]. PloS One，2015，10(7)：0130140.

[18] ZHOU B，KHOSLA A，LAPEDRIZA A，et al. Learning deep features for discriminative localization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition，2016：2921-2929.

[19] SHRIKUMAR A，GREENSIDE P，KUNDAJE A. Learning important features through propagating activation differences[C]//International Conference On Machine Learning，PMLR，2017：3145-3153.

[20] RIBEIRO M T，SINGH S，GUESTRIN C. " Why should i trust you?" Explaining the predictions of any classifier[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining，2016：1135-1144.

[21] LUNDBERG S M，LEE S I. A unified approach to

interpreting model predictions[C]// Advances in Neural Information Processing Systems，2017：30.

[22] MONGA V，LI Y，ELDAR Y C. Algorithm unrolling：Interpretable，efficient deep learning for signal and image processing[J]. IEEE Signal Processing Magazine，2021，38(2)：18-44.

[23] WRIGHT J，MA Y，MAIRAL J，et al. Sparse representation for computer vision and pattern recognition[J]. Proceedings of the IEEE，2010，98(6)：1031-1044.

[24] GREGOR K，LECUN Y. Learning fast approximations of sparse coding[C]//Proceedings of the 27th International Conference on International Conference on Machine Learning，2010：399-406.

[25] SUN J，LI H，XU Z. Deep ADMM-Net for compressive sensing MRI[C]//Advances in Neural Information Processing Systems，2016：29.

[26] LEFKIMMIATIS S. Universal denoising networks：a novel CNN architecture for image denoising[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition，2018：3204-3213.

[27] DARDIKMAN-YOFFE G，ELDAR Y C. Learned SPARCOM：Unfolded deep super-resolution microscopy[J]. Optics Express，2020，28(19)：27736-27763.

[28] PAPYAN V，ROMANO Y，ELAD M. Convolutional neural networks analyzed via convolutional sparse coding[J]. The Journal of Machine Learning Research，2017，18(1)：2887-2938.

[29] SULAM J，ABERDAM A，BECK A，et al. On multi-layer basis pursuit，efficient algorithms and convolutional neural networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2019，42(8)：1968-1980.

[30] BECK A. First-order methods in optimization[M]. Philadelphia：Society for Industrial and Applied Mathematics，2017.

[31] PARIKH N，BOYD S. Proximal algorithms[EB/OL]. https://stanford.edu/~boyd/papers/pdf/prox_slides.pdf.

[32] MOHRI M，ROSTAMIZADEH A，TALWALKAR A. Foundations of machine learning[M]. Massachusetts：MIT Press，2018.

[33] AKCAY S，ATAPOUR-ABARGHOUEI A，BRECKON T P. Ganomaly：Semi-supervised anomaly detection via adversarial training[C]//Computer Vision–ACCV 2018：14th Asian Conference on Computer Vision，Perth，Australia，December 2-6，2018，Revised Selected Papers，Part III 14. Springer International Publishing，2019：622-637.

[34] 许鹭. 航空发动机中介轴承故障综合诊断技术研究[D]. 沈阳：沈阳航空航天大学，2018.
XU Lu. Research on comprehensive fault diagnosis technology of aero-engine intermediate bearing[D]. Shenyang：Shenyang Aerospace University，2018.

[35] RAVANELLI M，BENGIO Y. Interpretable convolutional filters with sincnet[J/OL]. [2024-03-12]. https：//arxiv.org/abs/1811.09725.

作者简介：王诗彬，男，1985 年出生，博士，教授，博士研究生导师。主要研究方向为航空发动机与直升机故障诊断和健康管理。

E-mail：wangshibin2008@xjtu.edu.cn

陈雪峰(通信作者)，男，1975 年出生，博士，教授，博士研究生导师。主要研究方向为复杂机电装备动态特性分析与可靠性测试分析、故障诊断与健康管理等。

E-mail：chenxf@xjtu.edu.cn