

基于深度 SVDD 的发动机外涵静子叶片故障预警

史昊天, 蔡景, 程冲

(南京航空航天大学民航学院, 南京 211106)

摘要: 外涵静子叶片是大涵道比涡扇发动机气路的核心部件之一, 外涵静子脱出是一种较为严重的故障模式, 此故障可能会导致飞机或其他发动机部件损伤, 进而造成灾难性事故。对外涵静子叶片脱出故障的预警是一项重要的工作。但因其早期特征不明显, 现有的方法较难对此类故障进行有效的预警。因此, 针对该问题, 基于监控数据提出一种深度特征提取的支持向量数据域描述 (Support vector data description, SVDD) 的故障预警方法, 以实现对外涵静子叶片脱出故障的早期预警。首先, 采用基于发动机气路性能辨识的建模方法, 建立发动机特定性能参数的观测模型对气路参数进行深度特征提取, 以真实状态量与模型观测量的差值作为航空发动机是否发生故障的特征; 然后利用 SVDD 算法建立决策边界, 实现故障数据的自动划分, 决策边界生成的阈值可在故障发生之前的一定时间之内给出告警; 最后, 经过多次计算, 结果表明, 在故障早期直至故障发生的区间内, 表征其健康状态的性能参数都与观测量有较大的偏移, 表明了所选特征的有效性。使用数据增强方法生成故障仿真数据与真实数据进行对比验证, 预警时间比故障真实发生时间预警模型平均提前 3.14 h。

关键词: 故障预警; 外涵静子脱出; 系统辨识建模; 神经网络; 健康序列

中图分类号: V235.13

文献标志码: A

文章编号: 1005-2615(2024)05-0939-11

Fault Warning of Engine Fan Outlet Guide Blades Based on Deep-SVDD

SHI Haotian, CAI Jing, CHENG Chong

(College of Civil Aviation, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China)

Abstract: Fan blades are ones of the core components of the gas path of a high bypass ratio turbofan engine, and fan outlet guide vane detachment is a severe failure mode. This failure could potentially damage the aircraft or other engine components, leading to catastrophic accidents. Therefore, early warning of fan outlet guide vane detachment has become an important task. However, due to the subtle early features of this type of failure, existing methods struggle to effectively warn against it. Therefore, to address this issue, a failure warning method based on deep feature extraction and support vector data description (SVDD) is proposed using monitoring data, aiming to achieve early warning of fan outlet guide vane detachment. First, a modeling method based on engine gas path performance identification is used to establish an observation model of specific engine performance parameters for deep feature extraction. The difference between the real state quantity and the model observation quantity is used as the feature of whether the aero engine has a failure. Second, the SVDD algorithm is used to establish a decision boundary, realizing the automatic division of failure data. The threshold generated by the decision boundary can provide an alarm within a certain time before the failure occurs. Finally, after multiple calculations, the results show that in the interval from the early stage of the failure to the occurrence of the failure, the performance parameters characterizing its health status have a large deviation from the observation quantity, indicating the effectiveness of the selected

基金项目: 民航安全能力建设基金(2021-104)。

收稿日期: 2024-03-03; **修订日期:** 2024-05-28

通信作者: 蔡景, 男, 副教授, 硕士生导师, E-mail: caijing@nuaa.edu.cn。

引用格式: 史昊天, 蔡景, 程冲. 基于深度 SVDD 的发动机外涵静子叶片故障预警[J]. 南京航空航天大学学报, 2024, 56(5): 939-949. SHI Haotian, CAI Jing, CHENG Chong. Fault warning of engine fan outlet guide blades based on deep-SVDD[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2024, 56(5): 939-949.

features. The failure simulation data generated using data augmentation methods are compared and verified with real data. Compared with the actual time of the failure, the warning model realizes the warning of the failure on average 3.14 h in advance.

Key words: fault warning; fan outlet guide vane detachment; system identification modeling; neural networks; health index

大涵道比涡扇发动机具有耗油率低、起飞动力大、噪音低和迎风面积大等特点,被广泛运用到巨型、远程运输机与宽体客机之中。其动力有80%来源于外涵道,而在外涵道的各种部件中,外涵静子叶片(风扇出口导向叶片)在发动机工作过程中有提高气体压力、降低流速和改变流向的作用,是涡扇发动机最具有代表性的重要零件。涡扇发动机的性能与它有着密切的关系^[1-2]。

在涡扇发动机实际使用中,作为静子件的出口导向叶片发生故障的概率较小,但是由于设计或制造缺陷也会出现导致其断裂而脱出的事件,造成严重后果^[3]。针对外涵静子叶片脱出的问题,目前的研究主要是通过分析叶片裂纹与断裂面来改进设计方案。卜嘉利等^[4]运用荧光探伤检测、断口宏观分析、材质分析等方法,建立有限元应力模拟分析模型,对裂纹的性质及萌生原因进行细致分析,从而对航空发动机的设计与运用提出建议。李洋等^[5]通过对故障叶片进行外观检查、断口分析、表面形貌检查、截面金相检查、材质分析及断口区域成分分析,对叶片上缘板排气边转接区域的应力分布进行计算,确定了叶片裂纹性质和产生原因,并提出增加其距离的改进建议。García-Martínez等^[6]使用光学和扫描显微镜对叶片断口进行分析得出,疲劳现象是由叶片根部表面的点蚀引起的,并且随循环应力和温度梯度而发展。但是上述论断全部聚焦故障产生后的改进措施,对于正在使用的产品并不能防止故障发生。因此,在故障发生的早期对其故障特征进行识别,并对其所造成的影响进行评估隔离,不仅仅可以保障生产试验进度,而且对试验安全具有重要意义^[7]。由于早期故障在检查时较难发现,但其在数据层面表征出的差异可以被捕捉到。故拟选用基于用快速存储记录器(Quick access recorder, QAR)数据进行故障预警^[8]。

基于数据的故障预警方法十分多样,但是这些方法大多使用公开的仿真数据集,并致力于提升预警模型的泛用程度,造成了对于特定型号发动机的故障预警能力较弱的现象^[9]。然而,使用QAR数据进行故障预警存在着一定的困难:一方面试车试飞数据的故障标签难以获得;另一方面表征外涵静

子叶片的气路参数较少,振动信号参数也难以获得^[10]。因此,采用发动机气路辨识建模的方法建立健康状态的发动机观测模型,以此计算当前工况下健康的气路参数,用于与传感器采集参数作对比可以极大地提高故障预警的准确率。并且基于系统辨识的建模预警方法不需要精确的数学机理模型,只需要充分利用发动机领域的专家知识、经验和已有数据,便能对故障进行预警分析和诊断^[11-12]。基于数据的诊断方法在故障预警领域有着广泛的应用,不少方法已经取得了实质性的进展。卜旭东^[13]采用K-means聚类算法对试飞数据进行处理,然后使用反向传播神经网络(Back propagation neural network, BPNN)对涡扇发动机气路参数进行辨识,并将训练完成的模型应用于未参与训练的数据,以模型残差是否超过设定的阈值作为故障检测的依据,该算法能提前检测出故障。闫斌斌^[14]从气路性能诊断方法出发,研究基于混合模型的燃气轮机叶片故障预警及诊断中的若干关键问题。马浩祎等^[15]利用发动机仿真模型设计了微型涡喷发动机故障预警流程。毛浩英等^[16]利用深度学习方法建立了航空发动机故障预警模型,对20多种发动机故障进行预测提高了准确率。

基于系统辨识的发动机建模过程可看作一种回归过程,其本质是将输入的反应工作状态的变量通过线性或非线性模型输出为反应气路特征的性能变量。深度神经网络(Deep neural network, DNN)具有较强的非线性适应能力并且对于局部拟合有较强的特征学习能力,对于飞行工况进行辨识观测有着较好的性能^[17]。因此,选用DNN作为辨识观测的回归模型进行深度特征提取。并利用其局部最优特性训练其在特定状态下的拟合效果,训练拟合各种飞行工况下被监测参数的数值,并计算其残差,达到提取特征的目的。支持向量数据域描述(Support vector domain description, SVDD)作为一种基于距离的单值分类算法,常被应用于故障诊断领域^[18],在相对小规模训练集上能够取得良好的效果。这使得在无准确训练标签的情况下也能有效地进行故障检测。因此,使用SVDD方法,对多参数进行融合给出故障预警的阈值以达到故障自动预警的目的。

1 外涵静子叶片故障预警模型

外涵静子叶片位于中介机匣与风扇机匣之间。如图 1 所示,外涵静子叶片引导和整流进入外涵道的气流,使其能以最佳的角度和速度排出,从而最大化地产生推力。因此,当外涵静子叶片发生故障,气流的引导和整合发生异常,便会直接影响静子叶片后的静压,静子叶片前的进气静压,以及影响分流到内涵道的气流。压气机进气温度也会随之受到影响。

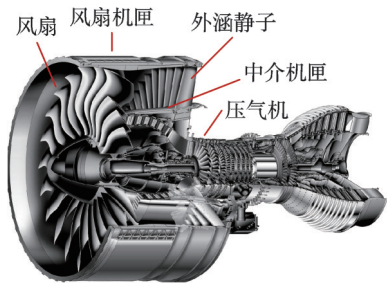


图 1 大涵道比的涡扇发动机示意图
Fig.1 Schematic diagram of a turbofan engine with external channels

由于发动机状态会反映在高低压轴参数上,所以用于发动机辨识建模,进行深度特征提取的参数有:低压轴转速(N_1)、高压轴转速(N_2)、高压压气机进口温度(T_1)、进气静压(p_1)、风扇静子后静压(p_2)等。使用基于系统辨识的气路热力学建模方法建立气路参数的观测模型。其中进气静压为进气道进气静压,风扇静子后静压为风扇出口导向叶片后的监测静压,高压压气机进口温度为低压压气机出口与高压压气机进口处的温度此截面风扇出口截面相近,风扇出口发生故障会对其造成影响。文献[13]使用的工作参数变量为:高度(H)、马赫数(Ma)、大气压强(p_0)、大气温度(T_0)、燃油流量(W)使用其对上述参数进行回归,具体建模映射关系如表 1 所示。

| 表 1 参数拟合关系 | |
|----------------------------------------|------------|
| Table 1 Parameter mapping relationship | |
| 输入参数 | 拟合参数 |
| H, Ma, T_0, p_0, W | N_1, N_2 |
| H, Ma, T_0, W | T_1 |
| H, Ma, p_0, W | p_1 |
| H, Ma, p_0, p_1, W | p_2 |

外涵静子叶片预警方法使用发动机气路特征观测模型作为故障特征的提取方法,使用单值分类算法给出自适应决策边界,将健康状态与故障状态

- 区分开。整体的故障预警流程整理步骤如下:
- 步骤 1 对航空发动机(健康状态)检测到的数据进行清洗整理变为适用于计算的数据;
 - 步骤 2 使用系统辨识法对航空发动机气动性能参数以及工况参数进行回归拟合训练出适用于辨识观测的数学模型;
 - 步骤 3 将航空发动机传感器监测的表征工作参数的数据输入到步骤 2 中建立的数学模型之中,生成此工作状态下航空发动机气路特征参数的预测值;
 - 步骤 4 将步骤 3 中生成的预测值与实际传感器监测到的数据做差,生成各气路特征参数的残差;
 - 步骤 5 使用多参数融合算法将上述生成的残差数据融合,生成表征健康状态的健康因子序列,并生成自适应退化阈值。

算法具体流程如图 2 所示。

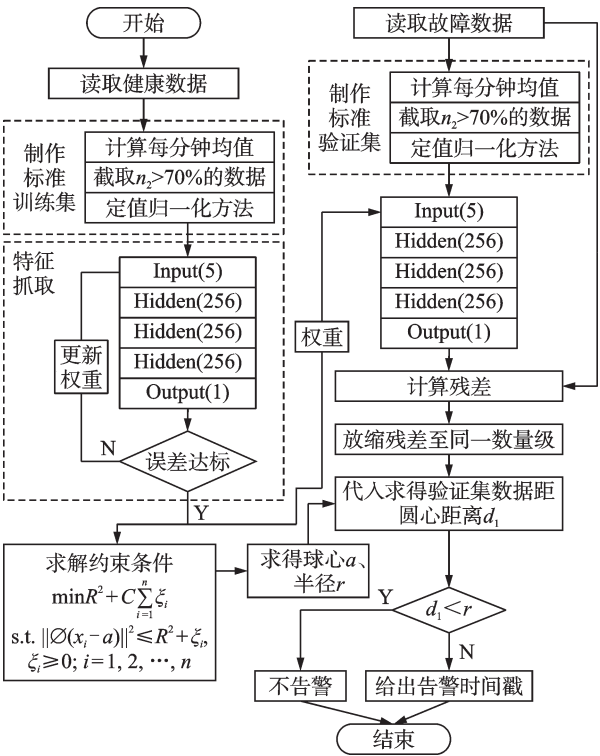


图 2 预警算法流程图
Fig.2 Warning algorithm flowchart

从图 2 中可以看出,整个预警流程分为两个部分,一部分为通过深度神经网络,建立系统辨识模型,此部分进行特征提取的方式为使用健康数据生成健康模型,将控制与工况参数输入特征提取模型获取气路参数真实量与模型量的残差,将残差作为 SVDD 算法的输入特征,最后利用 SVDD 得到的距离作为此时的健康序列的依据。

1.1 基于系统辨识的发动机气路特征观测模型

1.1.1 深度神经网络

DNN可以根据样本数据进行独立学习,从而更能表达输入和输出样本之间的映射关系,梯度下降法显著的降低了计算成本。它是一种多层前馈神经网络,并通过反向传播算法来调整网络的权重和偏置,从而最小化网络的输出误差。基于梯度下降优化的DNN的构建可以分为以下几个步骤。

(1)前向传播:给定一个输入向量 x 计算每个神经元的加权输入 z^l 和激活输出 a^l

$$z^l = W^l a^{(l-1)} + b^l \quad (1)$$

$$a^l = f(z^l) \quad (2)$$

式中: f 为激活函数; W^l 为每层的权重矩阵和偏置向量。

(2)反向传播:给定一个期望输出向量 y ,计算每一层的梯度由输出层开始向前递推。其中,输出层的梯度为

$$\frac{\partial E}{\partial z^L} = -(d - y) \odot f'(z^L) \quad (3)$$

隐藏层的梯度为

$$\frac{\partial E}{\partial z^l} = ((W^{(l+1)})^T \frac{\partial E}{\partial z^{(l+1)}}) \odot f'(z^l) \quad (4)$$

式中: E 为损失函数;“ \odot ”为哈达玛积(逐元素相乘); f' 为活函数的导数。

(3)梯度下降:根据每一层的梯度和激活值,使用梯度下降法更新参数。用梯度下降法更新参数时,需要指定一个学习率 η 后更新参数

$$\Delta W^L = -\eta \frac{\partial E}{\partial z^L} (a^{(L-1)})^T \quad (5)$$

$$\Delta b^L = -\eta \frac{\partial E}{\partial z^L} \quad (6)$$

式中: ΔW^L 和 Δb^L 为输出层权重和偏置的更新量;学习率 η 用于控制更新步长。

(4)参数寻优:重复以上步骤,直到网络的输出误差达到预设的阈值或者达到最大迭代次数。

1.1.2 数据归一化方法

QAR数据信息量大且采集频率较高,因此在进行模型训练之前需要对数据进行预处理。在涉及使用梯度下降的误差反向传播算法时,为了减少数据差异对于计算结果的影响需要考虑数据的归一化问题。在现有的解决方法中一般采用线性归一化法也称作最大最小值归一化,计算公式为

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (7)$$

式中: x 为待归一化的特征向量; $\min(x)$ 、 $\max(x)$ 为特征向量中的最大值与最小值; x' 为归一化之后的变量。然而在实际工程使用中,回归过程中输入

输出值存在着映射关系,实际模型中预测目标的数据包中最大最小值不同,致使使用式(7)进行归一化后的数据因为最大最小值存在着偏差,因此归一化之后的相对距离数据并不能代表真实的映射关系。为后续预测继续准确地使用该模型必须使用表征绝对距离的归一化方法。限 $\min(x)$ 、 $\max(x)$ 为定值,为方便计算以及避免数据溢出,将 $\max(x)$ 设为额定最大值的125%,将 $\min(x)$ 设为0,因此式(7)可变为

$$x' = \frac{x}{x^*} \quad (8)$$

式中 x^* 为额定最大值的125%。

1.1.3 特征提取模型训练评估与优化

为验证DNN作为观测模型的回归算法的可行性与优越性,使用回归分析中较为常用的智能算法:随机森林算法(Random forest, RF)、梯度提升回归树算法(Gradient boosting regression tree, GBRT)与BP(Back propagation)神经网络算法进行对比。为量化算法的优劣,使用回归分析中的均方根误差(Root mean square error, RMSE)、决定系数(R^2)、平均绝对误差(Mean absolute error, MAE)3个评价指标来量化算法性能,计算方法为

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (9)$$

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y} - y_i)^2} \quad (10)$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |(y_i - \hat{y}_i)| \quad (11)$$

式中: y_i 为实际值; \hat{y}_i 为模型预测值; \bar{y} 为真实值的平均值。

3种评价指标中, RMSE与MAE越小,说明模型预测值与真实值差距越小。决定系数越大,证明预测值与真实值之间的偏差越小,表示算法的性能越好。使用4种不同的回归算法对数据进行回归,得到3组数值具体结果如表2所示。由表2可以看出,5组性能参数回归中DNN算法的RMSE与MAE最小,并且DNN算法的决定系数并没有表现出很低的水平。因此认为DNN算法较为适合观测航空发动机气动性能参数。

但是,DNN训练的观测模型在全包线之下的效果并不好,通常DNN会更容易陷入局部最优。为保证建立模型的精确度,需要对进行计算的数据范围进行划分。在QAR数据中,慢车状态、飞行加速爬升等状态与高压轴转速有着密切的关系,通常是衡量飞机工作状态的一个重要指标,高压轴相对

| 表 2 4种算法的各项指标对比 | | | | | |
|--------------------------------------------------------------|-------|-------|-------|------|-------|
| Table 2 Comparison of parameters between the four algorithms | | | | | |
| 指标 | 参数 | RF | GBRT | BP | DNN |
| RMSE | N_1 | 7.2 | 7 | 0.88 | 0.91 |
| | N_2 | 1.4 | 1.4 | 1.6 | 1.12 |
| | p_1 | 6.6 | 7 | 2.1 | 1.1 |
| | p_2 | 3.3 | 3.5 | 2.2 | 1.2 |
| | T_1 | 2.05 | 1.9 | 1.8 | 1.5 |
| | | | | | |
| R^2 | N_1 | 0.77 | 0.79 | 0.85 | 0.96 |
| | N_2 | 0.89 | 0.899 | 0.88 | 0.97 |
| | p_1 | 0.905 | 0.895 | 0.83 | 0.976 |
| | p_2 | 0.96 | 0.956 | 0.82 | 0.99 |
| | T_1 | 0.87 | 0.89 | 0.76 | 0.95 |
| | | | | | |
| MAE | N_1 | 4.39 | 4.3 | 1.2 | 0.8 |
| | N_2 | 0.9 | 0.91 | 2.1 | 0.7 |
| | p_1 | 4.2 | 4.52 | 1.5 | 1.7 |
| | p_2 | 2 | 2.1 | 1.8 | 0.9 |
| | T_1 | 1.13 | 0.89 | 3 | 1.1 |

于转速这一单一参数,能更快地区分飞机的工作状态与飞行状态。因此,使用高压轴相对转速作为筛选航空发动机飞行数据阶段的指标,通过调整高压轴相对转速,建立回归模型训练测试结果的RMSE如图3所示。

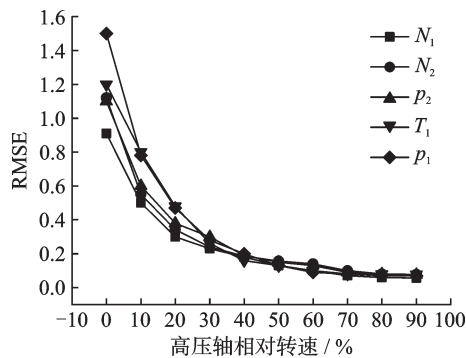


图 3 不同阶段各参数误差

Fig.3 Error of parameters at different stages

通过全阶段图像测试发现,高压轴相对转速为70% 以及其后的数据能使模型的RMSE稳定在0.1左右,因此认为此阶段以后观测模型的准确性有着较好的保证,由图4归一化后的高度与高压轴相对转速可以看出,当高压轴相对转速在70% 时发动机结束慢车状态,飞机开始起飞爬升。因此,在此区间的样本点可以认为能够较大的反应飞机的飞行状态,适用于建立飞行阶段的性能参数观测模型。

为直观表征模型准确度,截取一段飞行高度连续下降的数据,其下降过程中大气压强与大气温度等都在不断的变化,将其作为输入数据以测试观测模型,数据归一化形式如图5所示。

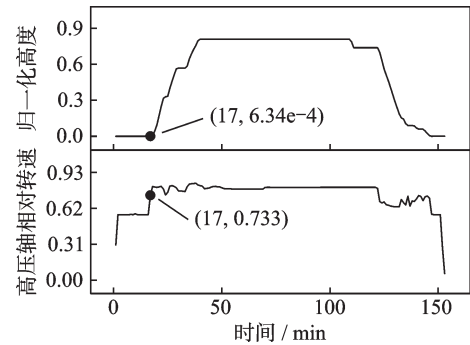


图 4 高压轴相对转速与高度关系

Fig.4 Relationships between high pressure shaft relative speed and height

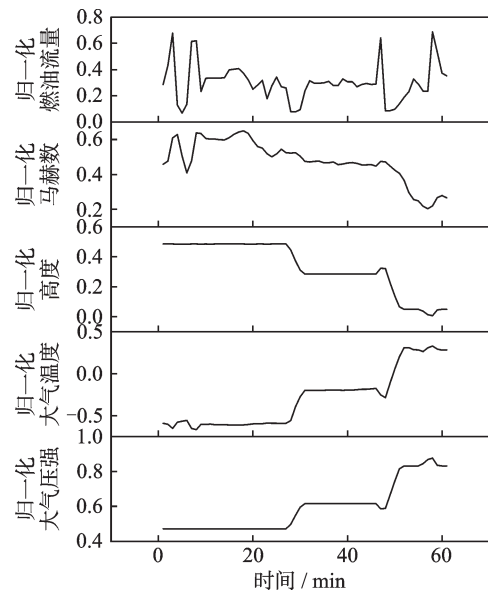


图 5 测试数据

Fig.5 Test data

使用DNN模型进行数据观测,以多层感知机作为DNN的训练模型,其网络结构示意图如图6所示。使用自适应矩估计(Adaptive moment estimation, AME)作为优化算法。使用Sigmoid函数作为激活函数。

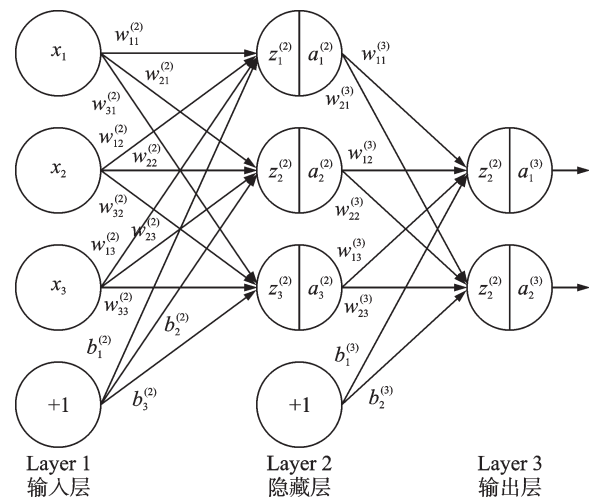


图 6 神经网络示意图

Fig.6 Schematic diagram of neural network

使用表1中的参数拟合关系分别构建拟合模型,使用网格搜索法调整超参数所得的收敛图像如图7所示。可以看出,当隐藏层层数等于4的时候,模型收敛性较差。通过改变隐藏层维度可以看出,当隐藏层在3层、维度在256的情况下模型收敛更快,效果更好。将数据输入模型中可得到拟合数据,如图8所示。

上述数据的均方根误差为0.09,并且可直观看出,辨识观测模型的拟合效果较好。因此认为此模型可以用于提取故障特征的观测模型。

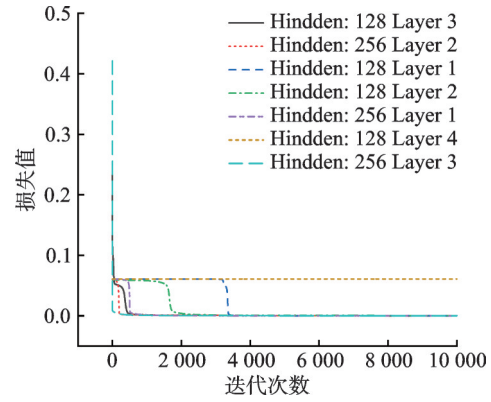


图7 损失下降图

Fig.7 Loss reduction chart

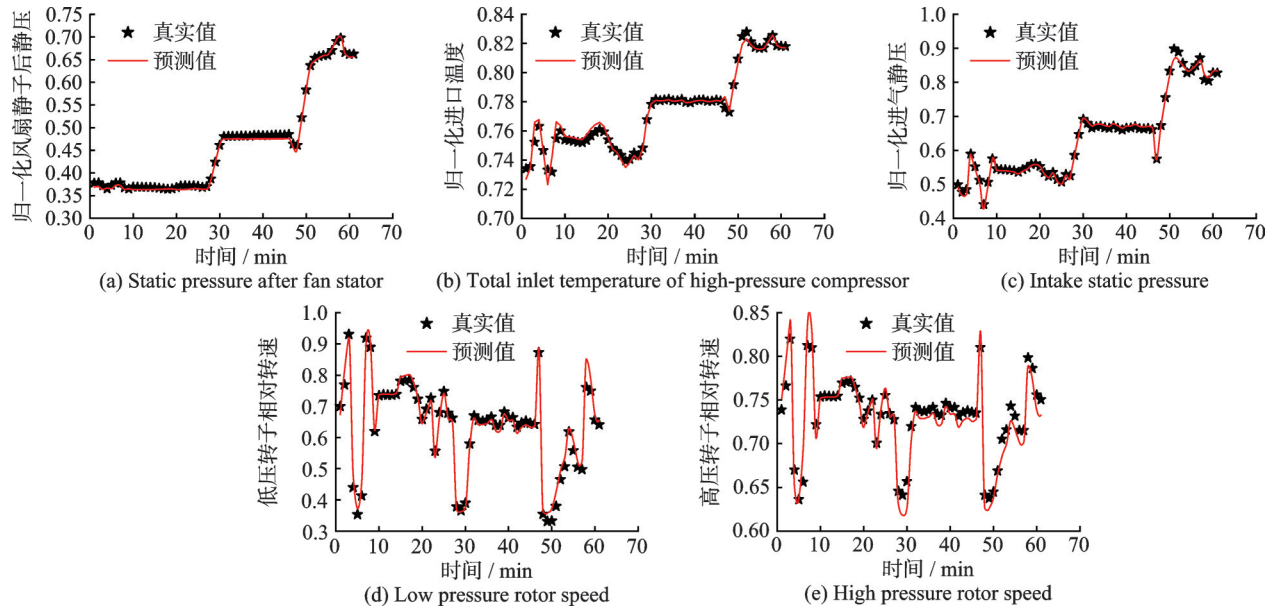


图8 测试集拟合效果

Fig.8 Fitting effects on the test set

1.2 外涵静子叶片故障预警方法

利用SVDD无监督的特性可以在无标签的情况下自动求解退化阈值,并且通过算法本身基于距离的特点能够将健康序列可视化。

1.2.1 支持向量数据域描述

支持向量数据域描述是一种基于单分类思想的算法。算法将数据点映射在高维特征空间中,构建一个超球体,求解一个将样本点包围的超球边界,并使该超球边界在半径尽可能小的情况下包含进尽可能多的目标样本,将非目标样本尽量排除在超球边界外。该方法通过控制半径、超球界面的形状以及样本数之间的关系达到平衡,从而将样本进行分类。

超球体的半径 r 由超球边界上的支持向量决定,并起到分开是否是目标样本的决策作用。根据到球心的距离,判断新的测试样本点在空间的位置,测试样本点距球心的距离小于半径,则认为测试样本点位于超球边界之内;反之,则认为其位于

超球边界之外,测试样本为异常样本。

假设一组数据 $x \in \mathbf{R}^{n \times d}$,其中 n 是样本个数, d 是特征维度。首先通过核函数(非线性变换函数) $\phi: x \rightarrow F$ 从原始空间映射到特征空间,在特征空间中寻找一个半径最小的超球体,尽可能地将所有的点包括在内。为了构造这样一个最小半径超球体,要解决以下优化问题

$$\min r^2 + C \sum_{i=1}^n \xi_i \quad (12)$$

$$\text{s.t.} \left\| \phi(x_i - a) \right\|^2 \leq r^2 + \xi_i \quad \xi_i \geq 0; i = 1, 2, \dots, n \quad (13)$$

式中: r 代表超球体半径; a 代表超球体的球心; ξ 为松弛因子; C 是一个权衡超球体半径的惩罚参数。结合拉格朗日乘子法,则原问题的对偶问题为

$$\min \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j (x_i, x_j) - \sum_{i=1}^n \alpha_i (x_i, x_i) \quad (14)$$

式中 α_i 为 x_i 的拉格朗日系数。拉格朗日系数满足 $0 \leq \alpha_i \leq C$,则称为支持向量。假设训练数据集中

属于支持向量的样本集合为 S_v , 则超球体的球心和半径的计算公式分别为

$$a = \sum_{i=1}^n \alpha_i \bigotimes (x_i)$$

(15)

$$r = \sqrt{(x_v, x_v) - 2 \sum_{i=1}^n \alpha_i (x_v, x_i) + \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j (x_i, x_j)}$$

(16)

式中 $x_v \in S_v$, (x_i, x_j) 为空间内积。测试样本 x_i 到超球体球心的距离为

$$d = \sqrt{(x_i, x_i) - 2 \sum_{i=1}^n \alpha_i (x_i, x_i) + \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j (x_i, x_j)}$$

(17)

若 $d < r$ 则说明测试样本在超球体边界之上或者内部, 属于正常样本; 反之则属于异常样本。

1.2.2 预警方法建模过程

航空发动机的退化过程是不可逆的, 必须经过人为恢复, 因此构建的健康因子应符合单调变化的趋势, 但又允许局部波动。使用观测残差作为输入特征使用, SVDD 算法进行降维, 使用点的距离作为构建因子的指标, 定义健康值为 H , 但由于健康值为从 1 开始下降, 则健康值的数学表达式为

$$H = 1 - d$$

(18)

为实现故障的自动预警, 在计算过程中首先使用健康数据与发动机观测模型之间的残差作为 SVDD 模型的训练数据, 建立 SVDD 模型, 并使用建立的模型对后面数据进行计算。得到数据点对于数据中心的偏移量, 以此作为健康指数曲线的模型, 并使用超球体半径作为健康指标的阈值。为优化决策边界, 建立性能更加优越的决策模型, 使用不同核函数进行互相对比, 具体使用的核函数有如下: 多项式核、高斯核、线性核、径向基函数(Radial basis function, RBF)核等。具体计算结果如表 3 所示。

表 3 不同核函数准确度

Table 3 Accuracy of different kernel functions

| 核函数 | 训练准确度/% | 测试准确度/% | 迭代次数 |
|-------|---------|---------|-------|
| 多项式核 | 85 | 82.3 | 500 |
| 高斯核 | 98.6 | 97.5 | 500 |
| 线性核 | 82.1 | 81.0 | 500 |
| RBF 核 | 96.6 | 93.2 | 500 |
| 双曲正切核 | 90.3 | 91.1 | 500 |
| 多项式核 | 86.2 | 83.6 | 1 000 |
| 高斯核 | 99.6 | 98.5 | 1 000 |
| 线性核 | 86.1 | 80.0 | 1 000 |
| RBF 核 | 97.9 | 95.3 | 1 000 |
| 双曲正切核 | 92.3 | 89.1 | 1 000 |

由表 3 可以看出, 使用基于高斯核函数的 SVDD 算法, 其准确度在训练与测试中表现均为优秀, 因此, 划分决策边界的 SVDD 模型的核函数使用基于高斯核函数的模型。

2 外涵静子叶片脱出预警实例及模型验证

2.1 实例计算

试验实例为某大涵道比涡扇发动机外涵静子脱出故障。数据提供了用于算法计算的控制与状态特征参量。其中控制参量如油门杆位置、放气活门开度、航空发动机运行状态字(慢车、起飞、巡航等), 航空发动机状态参量为航空发动机各截面的温度、压力、高低压轴转速、马赫数、燃油流量等。实例数据为某架次飞机外涵静子脱出故障及其前 10 次起飞循环, 数据总长度在 1 000 min。其中, 航行一次发动机的工作状态(归一化后)如图 9 所示。

为获取标准训练集与验证集, 对数据进行如下预处理。

步骤 1 将表 1 中的数据选出计算每 1 min 的平均值, 以达到去噪减少计算量的效果;

步骤 2 根据图 3 的计算结果提取高压转子相对转速在 70% 以上的数据点;

步骤 3 使用式(7,8)对步骤 1 提取到的数据进行归一化处理。

使用上述预处理过程对飞行数据进行清洗, 选取特定飞行状态区间之后, 剩余 931 min 的数据。已知的故障发生在数据第 900 min 左右。经过发动机辨识观测模型输出的测试与真实值的残差为深度神经网络提取到的特征, 其特征的意义为真实数据与发动机健康状态模型的偏差值, 偏差值可以代表发动机在此时的状态, 其残差结果如图 10 所示。

分析图 10 可得知, 样本数据在 500 min 之前波动较小, 参数残差基本在 0 附近。但随着故障特征的进一步凸显, 参数残差进一步变得不稳定, 最后当故障发生时, 相关参数的残差有着明显的突变。为进一步探寻退化规律, 实现故障的提前预警, 使用已训练好决策边界的 SVDD 模型来划分正常与异常数据。其中 SVDD 模型数据输入值为风扇静子后静压残差、低压轴转速残差、高压轴转速残差、进气压力残差、高压压气机进口温度残差, 并使用归一化方法进行归一化处理。通过 SVDD 的统计计算出一个阈值计算, 此阈值为健康与预警的重要指标。结果如图 11 所示。

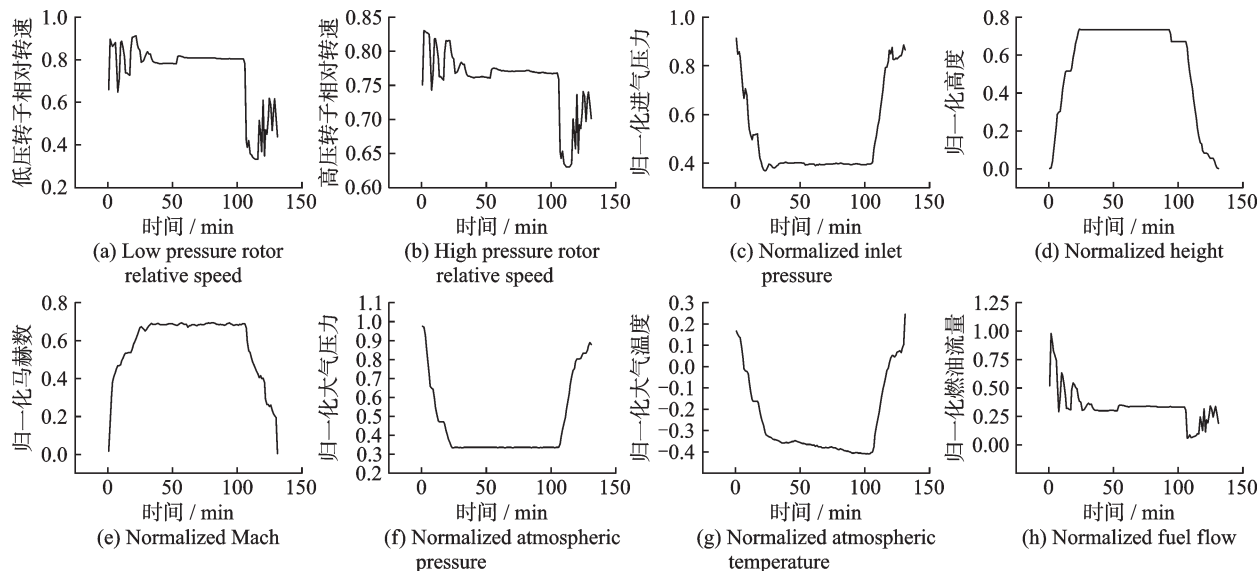


图9 某型涡扇发动机典型工作状态

Fig.9 Typical operating conditions of turbofan engines

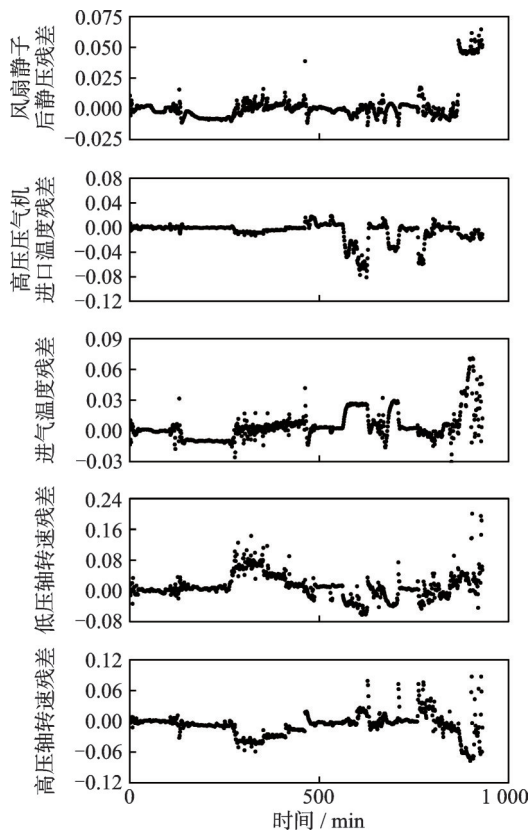


图10 测试与真实值之间的残差

Fig.10 Residual between test and true values

利用SVDD的决策阈值将数据分为健康数据与故障数据,并且每个数据的值代表着当前发动机的健康值。从图10可以看出越接近故障发生时间,发动机的健康值就越低下。为更好体现退化趋势,使用最小二乘法将离散的健康值拟合为一条曲线,效果如图12所示。

使用SVDD的决策阈值给出的预警点位在第694 min,故障点位于第900 min,本预警流程将故障预警时间提前了206 min(3.4 h)。

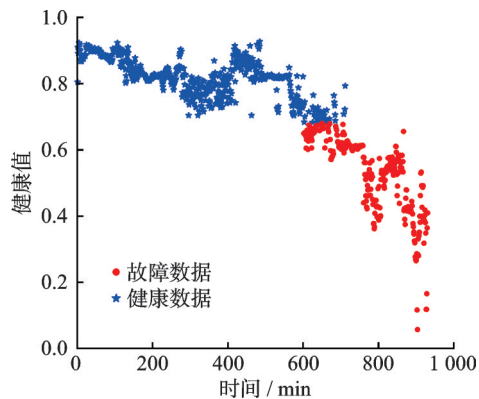


图11 健康与故障数据

Fig.11 Health and fault data

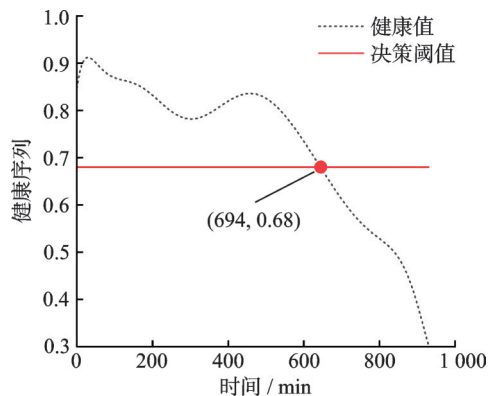


图12 拟合后的健康序列

Fig.12 Fitted health sequence

2.2 数据增强

由于故障产生于真实工况下的航空发动机,故障模式获取困难。为了弥补数据不足带来的问题。研究者们提出了小样本学习理论。其中,Yoon等^[19]提出TimeGAN(Time generative adversarial network)对小样本学习的数据增强方法。Deng等^[20]利用TimeGAN对病人是否患有高血糖

或者低血糖的预测问题进行数据增强,取得了较好的成果。因此使用TimeGAN方法来生成故障仿真数据,以验证预警流程的有效性。

2.2.1 TimeGAN

航空发动机故障数据一般由故障数据及其前序数据组成,其通过时间进行排列,属于时序数据。TimeGAN是在生成对抗网络(Generative adversarial network, GAN)的基础上诞生的,专门用于时间序列小样本问题。它有4个组成部分,通过训练3个损失函数来最小化生成数据与原始数据之间特征分布的差异。TimeGAN的4个组成部分分别为嵌入函数(Embedding function) E 、恢复函数(Recovery function)以及GAN网络中的序列鉴别器(Sequence discriminator) D 和序列生成器(Sequence generator) G 。

嵌入函数用于将原始时间序列数据映射到低维嵌入空间。恢复函数用于将数据从嵌入函数中的低维表示中恢复成原始时间序列数据,给定一个数据集 $\mathbf{X}=\{x_1, x_2, \dots, x_n\} \in \mathbf{R}^n$,则嵌入网络与恢复网络构建重构损失函数(\mathcal{L}_R)为

$$\mathcal{L}_R = E_{x \sim P(x)} \left(\left\| x - R(E(x)) \right\|_2^2 \right) \quad (19)$$

在时间序列数据下为了生成近似于目标分布的合成数据,给定一个随机向量 $\mathbf{Z} \in \mathbf{R}^n$,则对于序列生成器与序列鉴别器的损失函数(\mathcal{L}_{GAN})可定义为

$$\mathcal{L}_{GAN} = E_{x \sim P(x)} (\ln D(x)) + E_{z \sim P_z(z)} (\ln (1 - D(G(z)))) \quad (20)$$

对于时间序列而言,数据的时间维度与特征维度会对目标变量产生不同的影响,仅仅依靠GAN中的对抗训练与嵌入-恢复网络的非监督训练,会使生成的时间序列数据失真。因此,要求生成器在生成数据的同时,能够预测下一时间步的数据,其预测损失函数(\mathcal{L}_G)为

$$\mathcal{L}_G = E_{x \sim P(x)} \left(\left\| D(E(x)) - D(G(z)) \right\|_2^2 \right) \quad (21)$$

2.2.2 故障仿真

使用TimeGAN网络生成80批次数据,其中使用长短期记忆网络(Long short-term memory, LSTM)作为TimeGAN网络的生成器与判别器,每10批次取平均作为一组合成的数据,每组合成的数据为931 min,与原数据长度相同。图13为风扇静子后静压的8组仿真数据的部分数据,其中S1~S8分别代表着8组仿真数据。

为更直观表现仿真数据的特征,使用 t 分布邻域嵌入(t -distributed stochastic neighbor embedding, t -SNE)算法进行仿真数据的可视化方法,图

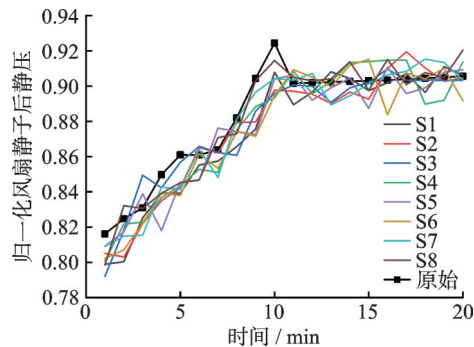


图13 部分仿真数据

Fig.13 Partial simulation data

14为仿真数据S1的 t -SNE可视化分析,其余仿真数据与此相仿在此不再赘述。

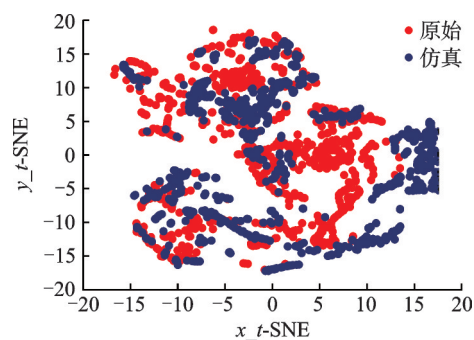


图14 t -SNE可视化

Fig.14 Visualization of t -SNE

由图14可以看出,仿真数据可以有效地保留原本数据的特征。因此,认为仿真数据可以用来验证预警算法流程。

2.3 对比实验验证

为验证模型的性能,将下列深度异常检测方法:深度自编码孤立森林(Deep autoencoding isolation forest, DAE-IF)、谱残差卷积神经网络(Spectral residual CNN, SR-CNN)^[21]、深度自编码高斯混合模型(Deep autoencoding gaussian mixture model, DAGMM)^[22]、长短期记忆网络编解码器(Long short-term memory network encoder-decoder, LSTM-ED)、深度对比单分类异常检测模型(Contrastive one-class anomaly, COCA)^[23]、对比预测编码异常检测(Contrastive predictive coding anomaly detection, CPC-AD)^[24]与所提出的带残差过程的深度SVDD(Deep support vector data description, DSVDD)算法进行对比验证。由于研究对象为真实发生的故障,因此数据的标签属性并不明显,只能明确到故障发生的时刻(第900 min),并不能将标签明确至退化发生时刻。因此为更好地描述算法的性能,采用误报率(False positive rate, FPR)指标与AUC指标作为评价算法优劣的指标,同时也对比各算法提前预警的能力。

$$FPR = \frac{FP}{FP + TN} \quad (22)$$

$$TPR = \frac{TP}{TP + FN} \quad (23)$$

$$AUC = \int_0^1 TPR(t) dt \quad (24)$$

通过计算仿真数据与原始数据,可以得到上述指标的平均值,所得具体数据如表4所示。

表4 7种算法性能对比图

Table 4 Performance comparison of seven algorithms

| 算法名称 | FPR/% | AUC | 提前时间/min |
|---------|-------|------------|----------|
| DAE-IF | 30.3 | 0.79±1.77 | 358.0 |
| SR-CNN | 12.73 | 0.848±0.38 | 85.8 |
| DAGMM | 4.24 | 0.849±0.22 | 156.0 |
| LSTM-ED | 18.18 | 0.891±0.69 | 15.4 |
| COCA | 7.89 | 0.92±0.43 | 108.9 |
| CPC-AD | 34.54 | 0.77±0.87 | 127.5 |
| DSVDD | 2.4 | 0.887±0.13 | 188.75 |

经过表4的对比可以发现DSVDD方法的误报率为几种算法中最低的,并且AUC指标表现良好。

图15为各算法预警时间的箱线图。可以看出DAE-IF有着较大的误差,并且其误报率也非常高,因此此算法不适合进行外涵静子叶片脱落的预警。结合图15与表4可以看出,DSVDD算法有着较低的误报率,并且预警时间的1.5倍四分位距(Interquartile range, IQR)也比较小,与之性能较近的DAGMM的预警范围与之相近。因此,DSVDD有着较好的预警能力,并且平均将预警时间提升至故障发生前188.75 min (3.14 h),标准差为21.2 min。

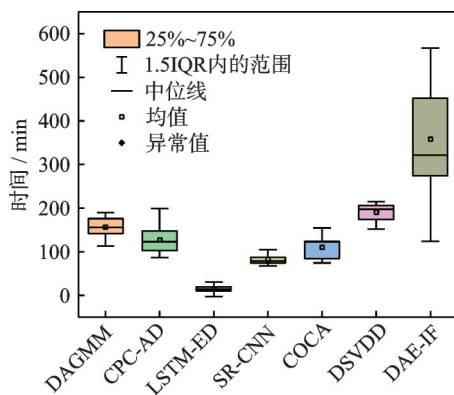


图15 算法预警提前时间对比

Fig.15 Comparison of early warning time for seven algorithms

3 结 论

本文针对航空涡扇发动机外涵静子叶片故障

前期特征不明显,故障预警难等问题,提出了一种基于系统辨识的故障预警方法,使用DNN算法实现特征参数的深度提取,最后利用SVDD算法自动计算决策边界给出预警阈值,达到故障提前预警的目的。具体结论如下:

(1)在外涵静子叶片失效脱出前期,由于疲劳、腐蚀、磨损等因素会使叶片故障早期发生振动或者形变从而体现在发动机气动性能上,其会对临近部件的气动性能产生一定的影响,特别是对风扇静子后静压这一参数影响明显。

(2)提出了一种提取风扇导向叶片断裂脱出的早期预警方法,并使用实际飞行数据与故障仿真数据进行验证,并通过多方法对比发现故障预警方法能将预警平均提前3.14个小时。

参考文献:

- [1] 党永杰,邵伏永,王春雪,等.弯扭静叶对跨声速单级风扇性能影响的数值研究[J].战术导弹技术,2014(3):79-83.
DANG Yongjie, SHAO Fuyong, WANG Chunxue, et al. Numerical research on the bowed-twisted blade effect on the transonic single fan performance[J]. Tactical Missile Technology, 2014(3): 79-83.
- [2] 王前,胡骏,李亮,等.弯曲静子叶片对叶尖射流扩稳效果的影响[J].航空动力学报,2015,30(12):2970-2975.
WANG Qian, HU Jun, LI Liang, et al. Effect of bowed stator blade on stability enhancement by tip injection[J]. Journal of Aerospace Power, 2015, 30(12): 2970-2975.
- [3] 黄超伟,马乐.航空发动机涡轮叶片故障诊断分析方法[J].内燃机与配件,2021(22):125-126.
HUANG Chaowei, MA Le. Analysis method of fault diagnosis of turbine blades for aero engines[J]. Internal Combustion Engine & Parts, 2021(22): 125-126.
- [4] 卜嘉利,高志坤,牛建坤,等.航空发动机风扇静子叶片裂纹失效分析[J].航空发动机,2021,47(6):91-95.
BU Jiali, GAO Zhikun, NIU Jiankun, et al. Crack failure analysis of a fan stator vane[J]. Aeroengine, 2021, 47(6): 91-95.
- [5] 李洋,邱丰,佟文伟.发动机风扇静子叶片裂纹失效分析[J].失效分析与预防,2019,14(6):401-405.
LI Yang, QIU Feng, TONG Wenwei. Cracking analysis of stator blade of aero-engine fan[J]. Failure Analysis and Prevention, 2019, 14(6): 401-405.
- [6] GARCÍA-MARTÍNEZ M, DEL HOYO GORDILLO J C, VALLES GONZÁLEZ M P, et al. Failure study of an aircraft engine high pressure turbine

- (HPT) first stage blade[J]. *Engineering Failure Analysis*, 2023, 149: 107251.
- [7] HANACHI H, MECHEFSKE C, LIU J, et al. Performance-based gas turbine health monitoring, diagnostics, and prognostics: A survey[J]. *IEEE Transactions on Reliability*, 2018, 67(3): 1340-1363.
- [8] SHUR M, STRELETS M, TRAVIN A, et al. Unsteady simulations of a fan/outlet-guide-vane system: Aerodynamics and turbulence[J]. *AIAA Journal*, 2018, 56(6): 2283-2297.
- [9] 黄金泉, 王启航, 鲁峰. 航空发动机气路故障诊断研究现状与展望[J]. *南京航空航天大学学报*, 2020, 52(4): 507-522.
- HUANG Jinquan, WANG Qihang, LU Feng. Research status and prospect of gas path fault diagnosis for aeroengine[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2020, 52(4): 507-522.
- [10] GUO C C, SUN Y C, SU S Y, et al. Risk assessment method for controlled flight into terrain of airlines based on QAR data[J]. *Aircraft Engineering and Aerospace Technology*, 2023, 95(8): 1184-1193.
- [11] BAI M L, LIU J F, LONG Z H, et al. A comparative study on class-imbalanced gas turbine fault diagnosis[J]. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 2023, 237(3): 672-700.
- [12] NASHED M S, RENNO J, MOHAMED M S, et al. Gas turbine failure classification using acoustic emissions with wavelet analysis and deep learning[J]. *Expert Systems with Applications*, 2023, 232: 120684.
- [13] 卜旭东. 基于试飞数据的涡扇发动机模型辨识及故障检测研究[J]. *工程与试验*, 2022, 62(2): 10-12, 130.
- BU Xudong. Research on model identification and fault detection of turbofan engine based on flight test data[J]. *Engineering & Test*, 2022, 62(2): 10-12, 130.
- [14] 闫斌斌. 基于气路性能混合模型的燃气轮机叶片故障预警及诊断方法研究[D]. 北京: 北京化工大学, 2021.
- YAN Binbin. Research on gas turbine blade fault early warning and diagnosis method based on gas path performance mixed model[D]. Beijing: Beijing University of Chemical Technology, 2021.
- [15] 马浩伟, 蔡景, 杨天策. 微型涡喷发动机喘振故障的预警技术研究[J]. *中国安全科学学报*, 2023, 33(S1): 209-214.
- MA Haoyi, CAI Jing, YANG Tiance. Study on early warning technology of surge fault of micro turbojet engine[J]. *China Safety Science Journal*, 2023, 33(S1): 209-214.
- [16] 毛浩英, 孙有朝, 李龙彪, 等. 基于改进DRSN的航空发动机故障风险预警模型[J]. *航空动力学报*, 2024, 39(2): 138-148.
- MAO Haoying, SUN Youchao, LI Longbiao, et al. Aeroengine fault risk early warning model based on improved DRSN[J]. *Journal of Aerospace Power*, 2024, 39(2): 138-148.
- [17] 李扬, 王逸斌, 朱春玲, 等. 基于遗传算法优化神经网络的结冰环境中MVD和LWC预测[J]. *南京航空航天大学学报*, 2023, 55(2): 282-290.
- LI Yang, WANG Yibin, ZHU Chunling, et al. Prediction of MVD and LWC in icing environment based on genetic-algorithm-optimized neural network[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2023, 55(2): 282-290.
- [18] ZHAO Y P, XIE Y L, YE Z F. A new dynamic radius SVDD for fault detection of aircraft engine[J]. *Engineering Applications of Artificial Intelligence*, 2021, 100: 104177.
- [19] YOON J, JARRETT D, SCHAAR M V D. Time-series generative adversarial networks[C]//*Proceedings of the 33rd International Conference on Neural Information Processing Systems*. New York, USA: ACM, 2019: 5508-5518.
- [20] DENG Y X, LU L, APONTE L, et al. Deep transfer learning and data augmentation improve glucose levels prediction in type 2 diabetes patients[J]. *NPJ Digital Medicine*, 2021, 4(1): 109.
- [21] REN H S, XU B X, WANG Y J, et al. Time-series anomaly detection service at microsoft[C]//*Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Anchorage, USA: ACM, 2019: 5508-5518.
- [22] BHATNAGAR A, KASSIANIK P, LIU Chenghao, et al. Merlion: A machine learning library for time series[EB/OL]. (2021-09-20). <https://doi.org/10.48550/arXiv.2109.09265>.
- [23] WANG R, LIU C W, MOU X D, et al. Deep contrastive one-class time series anomaly detection[C]//*Proceedings of the 2023 SIAM International Conference on Data Mining (SDM)*. Philadelphia, USA: Society for Industrial and Applied Mathematics, 2023: 694-702.
- [24] DE HAAN P, LÖWE S. Contrastive predictive coding for anomaly detection[EB/OL]. (2021-07-17). <https://doi.org/10.48550/arXiv.2107.07820>.