

基于 AQPSO-LSTM-BN 的 APU 故障诊断模型

高丹妮

(中国民航大学 电子信息与自动化学院,天津 300300)

摘要:针对飞机 APU 的常见故障,提出了一种基于量子粒子群优化的长短期记忆网络与批规范化相结合的飞机 APU 故障诊断模型。从 QAR 数据库中整理出需要的 APU 故障数据,将其进行归一化处理并分为训练集和测试集两部分,建立 CSV 文档数据库;对量子粒子群进行改进,使用改进后的量子粒子群对长短期记忆网络的隐含层单元数目进行寻优;将优化参数后的长短期记忆网络与批规范化相结合搭建网络模型,并在网络最顶层加入 Softmax 模型,生成 AQPSO-LSTM-BN 故障诊断模型,使用训练集对故障诊断模型进行训练,然后使用测试集进行实验。实验结果表明,该模型可以有效识别 APU 故障,与单一长短期记忆网络模型、支持向量机模型、循环神经网络模型和极限学习机模型相比,识别准确度有所提高。

关键词:长短期记忆网络;量子粒子群;自适应量子粒子群;故障诊断;飞机辅助动力装置;批规范化

中图分类号:TP206+.3

文献标识码:A

DOI: 10.19358/j.issn.2096-5133.2020.05.005

引用格式:高丹妮.基于 AQPSO-LSTM-BN 的 APU 故障诊断模型[J].信息技术与网络安全,2020,39(5):22-27.

Fault diagnosis model of APU based on AQPSO-LSTM-BN

Gao Danni

(College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Focusing on the common faults of aircraft APU, a fault diagnosis model of APU based on quantum particle swarm optimization and batch normalization is proposed. The model sorts out the required APU fault data from the QAR database, processes it into normalization, divides it into training set and test set, and establishes CSV document database. The improved quantum particle swarm is used to optimize the number of hidden layer units in the long and short term memory network. The network model is built by combining the optimized long and short term memory network with batch normalization, and at the top of the network model, Softmax model is added to generate the AQPSO-LSTM-BN fault diagnosis model. The fault diagnosis model is trained with the training set, and then the test set is used for the experiment. The experimental results show that this model can effectively identify the fault of APU and has higher recognition accuracy than LSTM model, support vector machine model, recurrent neural network model and extreme learning machine model.

Key words: long short-term memory; quantum particle swarm optimization; adaptive quantum particle swarm optimization; fault diagnosis; aircraft auxiliary power plant; batch normalization

0 引言

辅助动力装置(Auxiliary Power Unit, APU)作为飞机的重要装置,不仅可以保证飞机安全启动,在飞机停在地面时,还为飞机供气、供电,保证客舱舒适性。因此,对飞机 APU 进行故障诊断研究显得尤为重要。

APU 故障发生时,排故人员会结合故障发生的现场和自身的相关经验、故障手册的规定等对故障情况进行诊断和排查,这需要相关人员具有大量的相关知识、大量的运行维护经验和复杂的推理判断

能力。在大量故障信息同时出现时,即使经验丰富的机务人员也不能很好地完成处理。为了及时、准确地判断出真正的故障源,使用科学算法对 APU 进行故障诊断成了非常重要的一个课题。目前对飞机 APU 进行的科学研究有:张彭^[1]等人对 APU 的发展和应用现状进行了深入分析;邱亚华^[2]对 B737-300 机型 APU 启动时的常见故障进行了分析和排除;施开动^[3]等人对 A319 型飞机 APU 的启动故障做了可靠性调查。这些研究对搭建飞机 APU 故障诊断模型提供了有力的理论支持。GORI-

NEVSKY D^[4]使用基于模型的方法对飞机 APU 进行了故障诊断;刘铁庚^[5]等利用数学建模和仿真对 APU 进行了故障诊断。基于模型和数学建模都需要获得精确的飞机 APU 系统数学模型,模型的精确度直接影响最终的结果,具有一定局限性,并且飞机 APU 系统完整且精确的数学描述很难获得。唐启东^[6]等使用深度置信网络实现了对 APU 故障的检测并针对于此做了应对策略。深度置信网络缺少对历史信息处理,与之相比,长短期记忆网络(Long Short Term Memory Networks, LSTM)更加适于处理飞机 APU 故障信息。

目前,应用于飞机 APU 故障诊断的人工智能算法较少,常用的支持向量机、神经网络和仿生智能等单一算法均各有优缺点。支持向量机思想简单能有效解决非线性分类问题,但对核函数的选择十分敏感,无法处理大规模样本数据。人工神经网络具有很强的自学习能力,且可以较快地寻优,但隐含层数目等参数难以确定。仿生智能算法有较强的鲁棒性,但容易陷入局部最优。所以本文使用多个算法相结合的方法,利用改进的量子粒子群(Quantum Particle Swarm Optimization, QPSO)算法对 LSTM 进行了参数优化,然后针对于深层网络的梯度问题,利用批规范化层(Batch Normalization, BN)改进了 LSTM 网络模型的结构,建立 APU 故障诊断模型。得到自适应量子粒子群优化长短期记忆网络并结合批规范化(AQPSO-LSTM-BN)的故障诊断模型,对 APU 进行故障诊断。

1 自适应量子粒子群

1.1 改进量子粒子群

为了提高传统粒子群算法的收敛性和搜索能力,量子粒子群(QPSO)算法被提出,以下是 QPSO 算法的执行过程:

(1) 设置 $t=0$,并在 M 维待解空间中对量子粒子群各粒子进行初始化,随机获得各粒子的空间位置 $X_i(0)$,同时将各粒子的个体最好位置初始化为 $P_i(0) = X_i(0)$ 。

(2) 按式(1)求量子粒子群粒子平均历史最优值 mbest。

$$\text{mbest} = \frac{1}{M} \sum_{i=1}^M P_i \quad (1)$$

其中, M 为种群大小, P_i 为粒子的个体最优位置。

(3) 为每个粒子执行步骤(4)~(7)。

(4) 计算粒子的适应值,对粒子的个体最优位置 pbest 进行更新。其更新方法为将 $X_i(t)$ 的适应值与 $P_i(t-1)$ 的适应值进行比较,如果 $X_i(t)$ 的适应值优于 $P_i(t-1)$ 的适应值,即 $f[X_i(t)] < f[P_i(t-1)]$,则置 $P_i(t) = X_i(t)$,否则 $P_i(t) = P_i(t-1)$ 。

(5) 对于粒子 i ,比较其 $P_i(t)$ 与当前全局最优位置的 $P_g(t-1)$ 的适应值,若 $P_i(t)$ 的适应值优于 $P_g(t-1)$ 的适应值,即 $f[P_i(t)] < f[P_g(t-1)]$,则置 $P_g(t) = P_i(t)$,否则 $P_g(t) = P_g(t-1)$ 。

(6) 得到粒子群的各个粒子的随机位置。

(7) 按式(2)更新粒子的最新位置。

$$X_{i,j}(t+1) = Q_{i,j}(t) \pm \alpha \cdot |\text{mbest} - X_{i,j}(t)| \cdot \ln\left(\frac{1}{u(t)}\right) \quad (2)$$

$$Q_{i,j}(t) = \mu \cdot P_{i,j}(t) + (1 - \mu) \cdot P_{g,j}(t) \quad (3)$$

其中, $P_{i,j}(t)$ 和 $P_{g,j}(t)$ 分别为粒子 j 的个体最优位置和全局最优值, μ 为 $[0,1]$ 之间的随机数, α 为收缩扩张系数, $u()$ 是在 $[0,1]$ 上产生的一个随机数。

● (8) 若满足迭代停止条件,则终止迭代;否则, $t=t+1$,并返回步骤(2)。

QPSO 算法的运行中, α 为唯一的参数,其选择至关重要。常用的参数 α 的控制方式是选择一个固定值,这样会导致鲁棒性较差;或者从 1.0 线性减小至 0.5,这种控制方法普遍可得到一个良好的效果。但自适应控制的参数 α 效果更好,所以接下来对参数 α 进行自适应改进。

选取误差函数为 $\Delta = f_i - f_g$,其中, f_i 为当前粒子适应度值, f_g 为全局最优粒子适应度值,误差函数 Δ 表示粒子离全局最优位置的远近程度。由误差函数可以看出,误差函数越小,当前粒子离全局最优位置越近,搜索空间越小。为了使 QPSO 可以避免早熟现象的发生,需要让接近全局最优值的点和远离全局最优值的点拥有不同的 α 值。对于接近全局最优值的点,为了使种群收敛,赋予其较小的 α 值;远离全局最优值的点很难被搜索到,为了让进化继续下去,搜索到更宽的范围,应该赋予较大的 α 值。根据以上理论,本文提出一种自适应方式来调节参数 α 的值,如式(4)所示:

$$\alpha = \begin{cases} 0.6 - \frac{0.5\Delta}{f_m - f_g}, & \Delta \leq f_m - f_g \\ 1, & \Delta > f_m - f_g \end{cases} \quad (4)$$

其中, f_m 为粒子平均适应度值。此时在 QPSO 算法

中,参数 α 可随适应度值的变化而进行调整,即自适应量子粒子群算法(AQPSO)。

1.2 性能测试

选取 Schaffer 函数(二维)、Griewank 函数进行测试,两个函数全局最优值为 $(0, \dots, 0)$, 全局极值为 0。选取平均最优和标准偏差值为参考,对 AQPSO 算法进行性能测试,结果如表 1、表 2 所示。

表 1 Schaffer 函数测试结果

种群大小	迭代次数	PSO		AQPSO	
		平均最优	标准差	平均最优	标准差
25	1800	3.95E-06	0.001 362	0.001 876	0.003 528
50	1 800	5.04E-10	4.17E-04	0.001 086	0.003 124
100	1 800	2.79E-05	4.13E-09	1.48E-04	1.03E-10

表 2 Griewank 函数测试结果

种群大小	维度	迭代次数	PSO		AQPSO	
			平均最优	标准差	平均最优	标准差
25	10	800	0.103 54	0.092 61	0.084 25	0.057 30
	20	1 200	0.049 14	0.051 28	0.021 04	0.019 86
	30	1 800	0.021 37	0.037 06	0.013 02	0.016 28
50	10	800	0.090 13	0.085 40	0.075 19	0.051 28
	20	1 200	0.035 50	0.031 27	0.013 69	0.016 44
	30	1 800	0.014 88	0.016 91	0.010 16	0.011 61
100	10	800	0.080 39	0.082 95	0.052 46	0.021 37
	20	1 200	0.039 70	0.036 82	0.012 68	0.014 39
	30	1 800	0.015 28	0.046 11	0.006 47	0.070 02

仿真测试结果表明,AQPSO 算法与传统 PSO 算法相比,能够显著地增强搜索能力,在测试函数上得到了更好的结果。

2 LSTM 与批规范化

2.1 LSTM 网络结构

LSTM 是德国计算机科学家 Schmidhuber 在 1997 年提出的一种时间循环神经网络,可以改善过多存储信息导致的梯度消失、梯度爆炸等问题。如图 1 所示,每个 LSTM 单元在普通输出 h_t 的基础上增加了一个输出,即一个记忆单元 C_t ,并引入了三个门:输入门、遗忘门和输出门,通过这三个门控制了增加或者减少数据的比例。

LSTM 记忆单元与各个门控制器更新如下:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{5}$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{6}$$

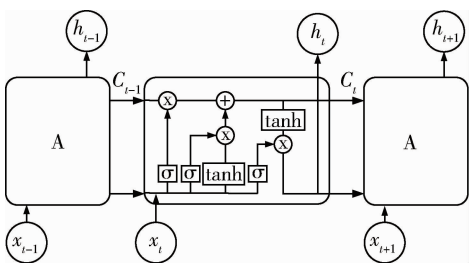


图 1 LSTM 结构图

$$c'_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1}) \tag{7}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot c'_t \tag{8}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{9}$$

$$h_t = o_t * \tanh C_t \tag{10}$$

式(5)、式(6)分别为输入门 i_t 和遗忘门 f_t 的更新公式,其中 b_i 和 b_f 是输入和遗忘的偏置矩阵, σ 为 sigmoid 函数, W 是相对应的权重矩阵。输入门表示了对输入和上一输出的接收量。遗忘门过滤了输入和上一输出的遗忘量。

式(7)为临时记忆单元 c'_t 的更新公式, $\tanh(\cdot)$ 为激活函数, W_{xc} 、 W_{hc} 分别为输入 x_t 和上一输出 h_{t-1} 的权重矩阵。

式(8)为最终的长期记忆单元 C_t 的更新公式,其是输入门 i_t 、遗忘门 f_t 、上一记忆单元 C_{t-1} 和临时记忆单元 c'_t 共同决定的结果。

式(9)是输出门更新公式,其中 W_{xo} 和 W_{ho} 是对应的权重矩阵。输出门的功能是对繁琐记忆信息的过滤。LSTM 单元最终输出如式(10)所示。

2.2 批规范化

BN 可以加快网络收敛速度,提升训练稳定性,且基本解决了深度网络难以训练的问题。多层神经网络很容易导致梯度问题,虽然 LSTM 中复杂的门结构可以一定程度上减少梯度消失的出现,但多层 LSTM 网络仍然会存在一些梯度问题,这时可以加入 BN 层对信号进行处理。BN 层首先对输入进行白化预处理:

$$\hat{x} = \frac{x - E(x)}{\sqrt{\text{Var}(x) + \varepsilon}} \tag{11}$$

其中, $E(x)$ 指其中一批输入 x 的平均值; $\text{Var}(x)$ 为该批次数据的标准差; ε 是保证分母不为零的微小正数。对于深层神经网络,在隐含层之后添加一层 BN,可抑制网络的非线性表达,降低网络的表达能力。为了保障网络的非线性获得,加入“比例及平移”操作,即:

$$y = \lambda \hat{x} + \beta = \text{BN}_{\lambda, \beta}(x) \quad (12)$$

其中, λ 是尺度因子, β 是偏移因子, 这两个参数通过训练获得。在这两个参数影响下, $\hat{x}$ 的图像从标准正态分布左移或者右移了一些, 使得非线性函数的值从正中心周围的线性区往非线性区做了细微的调整, 防止 BN 的“矫枉过正”。在极端条件下 $\lambda = \sqrt{\text{Var}(x) + \varepsilon}$, $\beta = E(x)$, 此刻 BN 层输出的数据和输入数据相同。所以通过 BN 层的信号可以更好地在 LSTM 中完成训练, 增加了模型的稳定性和运行速度。

3 搭建 AQPSO-LSTM-BN 故障诊断模型

3.1 网络搭建

基于 AQPSO-LSTM-BN 的故障诊断模型以 LSTM 为基本的网络单元, 使用 AQPSO 算法对 LSTM 隐含层节点进行优化, 并将 BN 层应用在 LSTM 隐含层上, BN 层对 LSTM 输出信息进行归一化并进行变换重构等修正处理, 改变了训练梯度, 使学习收敛速度加快, 大大加快了模型整体训练速度。顶层加入 Softmax 层对故障类别进行识别。AQPSO-LSTM-BN 网络结构如图 2 所示, 由输入层、加了 BN 的 LSTM 隐含层、一层全连接层和最顶层的 Softmax 层构成。

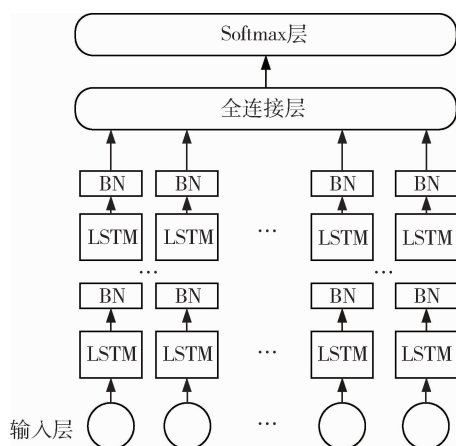


图2 AQPSO-LSTM-BN 网络结构

3.2 AQPSO-LSTM-BN 故障诊断模型的搭建

首先对从飞机 QAR 采集的故障数据进行归一化以及时间序列化预处理, 将预处理后的数据分为训练集和测试集。搭建故障诊断模型步骤如下:

(1) 按照 3.1 节搭建网络模型。选择 LSTM 隐含层, 每层 LSTM 上加入 BN 层, 将每个隐层的输出结果进行归一化处理再进入下一个隐层, 一层全连

接层, 顶层为 Softmax 分类层。

(2) 数据处理。对收集到的数据进行特征选择等预处理, 将处理后的数据分为训练集、测试集。

(3) 参数优化。选择 AQPSO 来对隐含层节点数进行寻优, AQPSO 优化 LSTM 过程如下:

①将 LSTM 中隐含层神经元作为优化对象, 并初始化粒子位置和速度。

②将粒子群各粒子改为量子态下模式。

③根据式(1)确定粒子平均最优位置, 按改进参数 α 后的式(2)更新粒子位置。通过比较目标函数, 选取全局最优值。重复此步骤直到满足终止条件。

④利用最优隐含层节点数构建 LSTM 网络。

在迭代学习中使用 Adam 算法优化网络参数, 最终使 Softmax 层的输出与真实数据标签的交叉熵损失逐渐减小。

(4) 当满足终止条件时停止训练, 否则返回步骤(1)。

通过训练集的数据训练得到准确的 AQPSO-LSTM-BN 故障诊断模型, 然后使用测试集的数据进行实验, 通过故障诊断模型给出分类结果, 并与真实类别作比较, 计算测试集上的正确率。具体流程如图 3 所示。

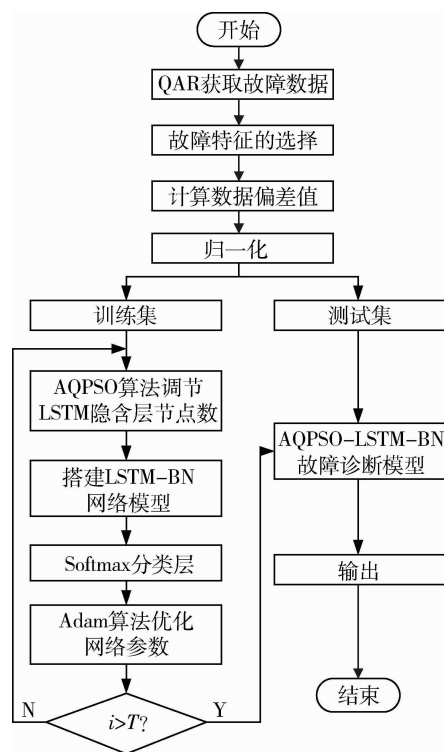


图3 AQPSO-LSTM-BN 故障检测流程图

4 实验与分析

4.1 实验数据

本文使用波音 787 飞机的 QAR 系统提供的飞机辅助动力装置 (APU) 的实时数据。人工选取适用于本实验的有效特征数据:排气温度 (EGT)、滑油温度 (OilTemp)、滑油量 (OilQty) 和转速 4 种故障参数,如表 3 所示。按照滑油故障、点火系统故障、发电机故障三类划分故障模式,使正常启动时输出为 0,滑油故障时输出为 1,点火系统故障时输出为 2,发电机故障时输出为 3。

表 3 故障特征选择

故障类别	故障种类	采样间隔/s
S1	滑油温度 (OilTemp)	1
S1	滑油量 (OilQty)	1
S2	排气温度 (EGT)	1
S3	转速	1

对故障参数求偏差值,将得到的偏差数据使用归一化进行预处理。归一化将数据缩放到[0,1]区间内,如式 (13) 所示:

$$y = \frac{x - \min}{\max - \min}$$

(13)

仿真实验中,keras LSTM 模型对数据形式有一定要求,通常为 3D tensor。因此,从 CSV 文档读入一组原始数据后,需要用 numpy.reshape 函数对原始数据进行重塑,以便 keras LSTM 读入。

4.2 模型参数对实验数据的影响

根据 AQPSO-LSTM-BN 故障检测器可知,在训练过程中对模型性能影响较大的参数有隐含层层数、隐含层节点数、训练次数 T 和学习率 lr 。

为了选择合适的隐层数量,分别选取隐层数为 1~5 的 5 种不同网络结构,并以自适应的量子粒子群 (AQPSO) 算法进行隐含层节点数选择,以 5 种层数的平均准确率为评判标准进行实验。经实验得出,隐层数为 3 和 4 时整体检测率相差不多,3 层与 4 层相比结构更简单,训练速度更快;隐层数为 5 时精度大幅度下降。综合考量则最合适的隐层数为 3,隐含层节点数为 50、50、24。

图 4 记录了训练次数对实验准确率的影响,描述了训练 2 000 次时,准确率的变化曲线 (假设 $lr = 0.006$)。

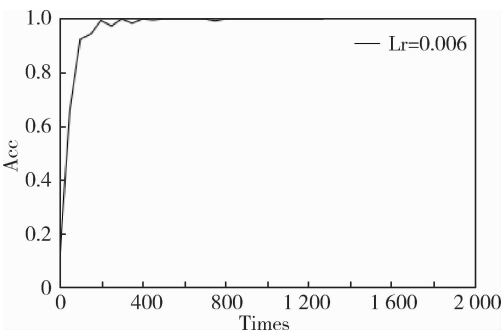


图 4 准确率随训练次数变化图

从图 4 可以看出,开始模型逐步收敛但前期不稳定产生震荡,这是由于选择了较大的学习率,导致参数更新太快。当训练次数 $T = 360$ 左右模型开始稳定。在训练次数选取时,若训练次数过少,会造成训练不足;若训练次数过多,则有可能会造成过拟合。这里将训练次数调整为 600 次。

图 5 为选择迭代次数为 600,学习率分别为 0.001、0.003、0.006 时的准确率变化曲线。

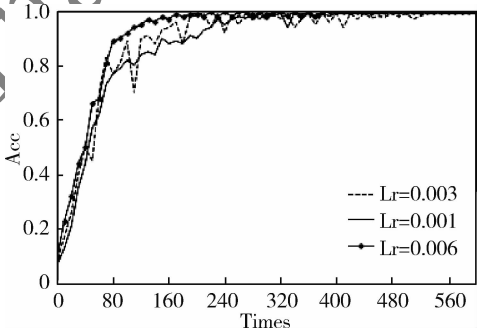


图 5 不同学习率准确率对比曲线

由图 5 可见,选择 $T = 600$,当 $lr = 0.001$ 时,准确率最终稳定在 99.3%;当 $lr = 0.003$ 时,准确率最终稳定在 99.1%,模型震荡最为剧烈且持续时间较长;当 $lr = 0.006$ 时,训练到 130 次时准确率开始收敛,收敛速度最快,且准确率较为稳定,最终收敛于 99.6%。由上述内容可知,选择学习率 $lr = 0.006$ 最为合适。

4.3 实验结果

设置隐含层节点数为 50、50、24,学习率 $lr = 0.006$,将 APU 故障含正常状态分为 0、1、2、3 共四类,选取 400 个参数向量作为测试集进行诊断,得到的诊断结果如图 6 所示。

图 6 中,横坐标 0~100 为正常数据,测试结果显示可以完全识别;100~200 为滑油故障,判错为正常 1 个;200~300 为点火系统故障,判错为滑油

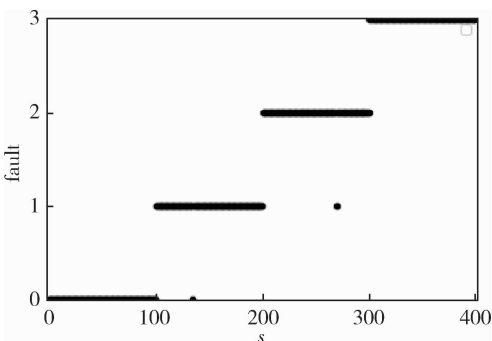


图6 AQPSo-LSTM-BN 诊断结果图

故障 1 个;300 ~ 400 为发电机故障,没有判错。为验证优化算法的合理性,将加了 BN 层的故障诊断模型与 AQPSo-LSTM 模型进行了对比,均选取训练次数 $T=600$ 次,对比结果如图 7 所示。

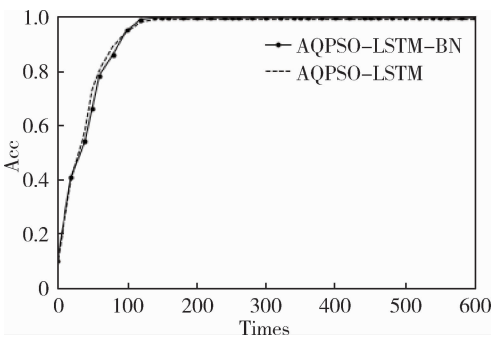


图7 AQPSo-LSTM 与 AQPSo-LSTM-BN 对比图

由图 7 可以看出,加 BN 层后的收敛速度与 AQPSo-LSTM 故障诊断模型相比有所提高,在训练次数 110 左右开始收敛,模型精确度也有所提高。为验证模型的准确性,将实验结果与传统的 LSTM 模型、AQPSo-LSTM 模型以及 SVM 模型、RNN 模型和 ELM 模型的准确率进行对比,结果如表 4 所示。

表 4 实验对比结果

方法	训练次数	准确率/%
AQPSo-LSTM-BN	600	99.60
AQPSo-LSTM	600	98.96
LSTM	800	97.68
SVM	800	95.65
RNN	800	94.37
ELM	800	92.55

由表 4 可以看出,针对于飞机 APU 故障数据, AQPSo-LSTM-BN 故障诊断模型与单纯对长短期记忆网络进行参数优化的故障诊断模型相比,准确率更高,且明显高于其他单一的故障诊断方法 (LSTM, SVM, RNN, ELM)。

5 结束语

常见的单一故障诊断方法各有优缺点,不能很好地进行诊断。本文提出了一种基于优化 LSTM 与 BN 相结合的故障诊断模型,首先针对 LSTM 参数难以确定的问题,使用改进的 QPSO 算法对 LSTM 隐含层节点数进行优化,然后针对深层网络梯度问题,使用 BN 对 LSTM 网络结构进行改进。实验在 Python 语言开发环境中编写程序并调用 TensorFlow 库函数完成。实验结果表明, AQPSo-LSTM-BN 模型可实现高精度的 APU 故障诊断,与普通 LSTM 模型、SVM 模型、RNN 模型、ELM 模型相比,在同等条件下, AQPSo-LSTM-BN 故障诊断模型的准确率更高,速度更快。本文提出的故障诊断模型更适用于处理时间序列问题,可以对飞机发动机、APU 等故障诊断提供一定参考。

参考文献

[1] 张彭. 飞机 APU 故障浅析和状态监控发展研究[J]. 科技资讯,2017,15(8):130-131.

[2] 邱亚华. B737-300 飞机 APU 启动常见故障的分析和排除[J]. 赤子(中旬),2014(14):178-179.

[3] 施开动,刘佳鸣. A319 飞机 APU 启动故障可靠性调查[J]. 中国民航飞行学院学报,2016,27(6):53-56.

[4] GORINEVSKY D, DITTMAR K, MYLARASWAMY D, et al. Model-based diagnostics for an aircraft auxiliary power unit[C]. Proceedings of the 2002 IEEE International Conference on Control Application. Glasgow, Scotland:IEEE, 2002,1:215-220.

[5] 刘铁庚. 某辅助动力装置试验故障分析[J]. 科学技术与工程,2012,12(33):9122-9125.

[6] 唐启东. APU 故障趋势预测及应对策略研究[D]. 天津:中国民航大学,2014.

(收稿日期:2020-04-01)

作者简介:

高丹妮(1995 -),通信作者,女,硕士,主要研究方向: APU 故障诊断与预测。E-mail:danygao_52@163.com。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所