

引用格式:戴婧睿,吴奇,仁和,等.基于深度置信网络的QAR飞行数据特征提取分析[J].电光与控制,2017,24(11):78-82. DAI J R, WU Q, REN H, et al. DBN based feature extraction for flight data of quick access recorder[J]. Electronics Optics & Control, 2017, 24(11):78-82.

## 基于深度置信网络的QAR飞行数据特征提取分析

戴婧睿<sup>1</sup>, 吴奇<sup>1</sup>, 仁和<sup>2</sup>, 裘旭益<sup>3</sup>

(1. 上海交通大学航空航天学院, 上海 200240; 2. 中国商用飞机有限责任公司, 上海 200241;  
3. 中国航空无线电电子研究所, 上海 200233)

**摘要:** 民航飞机的快速存取记录仪(QAR)在飞行过程中记录了大量的飞行参数, QAR数据是飞行安全评估的重要依据。针对QAR数据大样本、高维度的特点, 提出了一种有效的飞行数据特征提取的高效算法——DBN算法。DBN优势在于其能够摆脱对大量数据处理技术与专家经验的依赖而对飞行数据进行特征提取。在不同类别飞行数据集上进行仿真实验, 结果显示与主成分分析法(PCA)相比, 通过DBN提取的特征进行分类识别准确率更高。

**关键词:** 飞行数据; QAR; 数据分析; DBN模型; 特征提取

**中图分类号:** V271.1      **文献标志码:** A      **doi:**10.3969/j.issn.1671-637X.2017.11.016

## DBN Based Feature Extraction for Flight Data of Quick Access Recorder

DAI Jing-rui<sup>1</sup>, WU Qi<sup>1</sup>, REN He<sup>2</sup>, QIU Xu-yi<sup>3</sup>

(1. School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, China; 2. Commercial Aircraft Corporation of China, Shanghai 200241, China; 3. China Aeronautical Radio Electronics Research Institute, Shanghai 200233, China)

**Abstract:** A great number of flight parameters are recorded by the Quick Access Recorder (QAR) equipped on civil aircrafts. QAR data is an important criterion for flight safety assessment. Aiming at large-sample and high-dimension features of flight data from QAR, this paper proposes an effective feature extraction algorithm, Deep Belief Network (DBN) algorithm. The DBN algorithm can adaptively extract the features of flight data independent of data-processing technologies and expert experiences. Simulations of different types of flight data sets are carried out. The simulation results show that, compared with the PCA algorithm, the accuracy of classification and identification of features extracted by DBN model is higher.

**Key words:** flight data; QAR; data analysis; DBN model; feature extraction

### 0 引言

自从人类造出飞行器, 航空安全一直是人们关注的话题。为了飞行安全, 中国民航局强制要求国内航空公司必须安装快速存取记录仪(Quick Access Recorder, QAR), 可以记录飞行过程中一系列参数。通过QAR数据的分析, 可以实现飞行品质监控、飞行过程可视化仿真和事故调查<sup>[1]</sup>等功能。

QAR飞行数据具有高维度以及大样本的特征,

QAR数据的高维度是因为其记录了几百个飞行参数, 而样本大是因为QAR数据一次飞行可以连续记录100 h以上数据。所以针对QAR高维以及样本大的特征, 在飞行数据分析之前需要进行特征提取。

主成分分析(Principal Component Analysis, PCA)方法作为一种特征提取技术, 广泛应用于工业与科学领域<sup>[2]</sup>。文献[3]将PCA应用于无人机故障检测与诊断, 实现了无人机飞控系统传感器的故障检测和诊断; 文献[4]将PCA应用于QAR飞行数据特征提取, 并且通过分析, 相对于正常飞行数据, 研究异常飞行数据特点。PCA算法可以有效地找出数据中最“主要”的元素, 即可以提取对应于数据方差最大方向的变量(即主成分), 去除原始数据的噪音以及冗余。但是PCA对飞行数据进行特征提取的时候有其无法克服的缺点: 1) PCA将所有的

收稿日期: 2016-11-23

修回日期: 2017-01-21

基金项目: 国家自然科学基金(61671293); 上海浦江计划人才(15PJ1404300)

作者简介: 戴婧睿(1991—), 女, 陕西西安人, 硕士生, 研究方向为航空数据智能处理。

样本作为一个整体来对待,而数据方差较小的方向也会包含重要信息;2) PCA 是一个无监督的数据特征提取算法,当原始数据样本需要进行识别分类的时候,无法提取不同类别样本之间的差异性特征;3) PCA 为线性分析方法,而对于非线性飞行数据特征提取效果不佳。

为了解决以上问题,本文将引入深度置信网络(Deep Belief Network, DBN)作为飞行数据特征提取方法<sup>[5-6]</sup>。而深度置信网络具有强大的特征提取能力,可以解决数据特征提取、故障检测与分类等问题。DBN 特征提取技术相对于传统的 PCA 特征提取技术具有以下优点:1) DBN 可以自动从原始信号提取特征,减少了对专家经验与信号特征提取技术的依赖,减少了人工提取特征的不确定性;2) DBN 是一个深度网络模型,非常适合飞行数据的大样本、高维、非线性数据的特点;3) 由于 DBN 是深层的网络,通过多层的特征提取,可以弱化上层网络的错误特征;4) DBN 是一个有监督的特征提取方法,即可以提取不同类别数据的差异性特征,为分析不同类型的数据样本提供了有效的特征提取技术。由于 DBN 是一个深度结构,需要大量的样本进行训练来确定网络模型,本文将 DBN 应用于 QAR 飞行数据的特征提取,与 PCA 特征提取技术进行对比,并在不同类别飞行数据集上进行仿真实验,结果显示与 PCA 特征提取技术相比,通过 DBN 提取的特征进行分类识别准确率更高。

### 1 深度置信网络(DBN)

深度置信网络是成功地融合了深度学习与特征学习的多层神经网络。若干层无监督的限制玻尔兹曼机(Restricted Boltzmann Machine, RBM)与一层有监督的反向传播网络(Back-Propagation, BP)组成了深度学习中的一种网络模型,即 DBN,其结构见图 1。

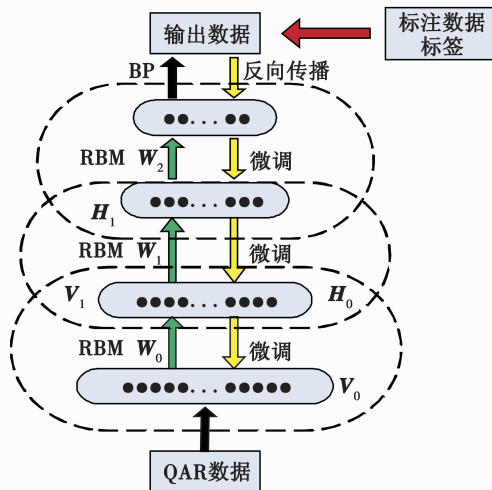


图 1 DBN 结构图  
Fig. 1 Structure of DBN

DBN 的训练过程分为两步:1) 分别由低层到高层无监督地训练每层 RBM,使特征向量映射到不同特征空间的时候能够保留尽可能多的原始特征信息<sup>[7]</sup>;2) 在最顶层利用 BP 网络模型对这个网络进行有监督的训练,微调整个网络模型,得到 DBN 的最优参数。DBN 模型通过将 RBM 结构与 BP 结构相结合提高了整个网络的性能。通过 RBM 网络可以初始化 DBN 模型的参数,但是 RBM 网络只能保证 DBN 模型各个层内的特征映射达到最优,却不能保证整个 DBN 模型内可以达到特征映射最优,即 RBM 学习到的是普遍的概念化特征。对于多类数据特征提取,RBM 网络不可以针对某一类数据提取最优特征,只可以提取普遍化特征,因此顶层需要有一个监督的学习对整体结构进行优化。所以在 DBN 模型的训练过程中,需要利用 BP 网络模型有监督地训练最顶层的模型,可以将 RBM 提取到的特征进行分类,同时可以微调整体 DBN 模型,将错误的信息返回到 RBM 网络模型中,微调 RBM 网络的参数,使 DBN 模型的参数达到最优,进而提取不同类别数据的差异性特征。

#### 1.1 限制玻尔兹曼机

限制玻尔兹曼机(RBM)是 DBN 模型的重要模块之一,也是玻尔兹曼机的一种特殊形式<sup>[8]</sup>。如图 2 所示,RBM 网络一般由两层组成,第一层为可见层,第二层为隐藏层。

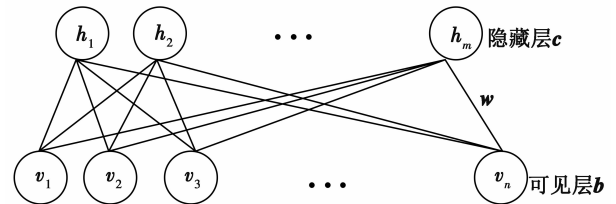


图 2 RBM 结构  
Fig. 2 Structure of RBM

RBM 网络中,下层是由  $n$  个可见节点组成的可视层,上层是由  $m$  个隐藏节点组成的隐藏层,也是特征提取层。 $\mathbf{b} = (b_1, b_2, \dots, b_n)$  为可视层的可见节点的偏移量, $\mathbf{c} = (c_1, c_2, \dots, c_m)$  为隐藏层的隐藏节点的偏移量, $\mathbf{w}_{n \times m}$  是可见节点与隐藏节点的权值矩阵。对于一些确定的网络,即  $\mathbf{v}, \mathbf{h}$  确定,RBM 的能量定义为

$$E(\mathbf{v}, \mathbf{h}) = - \sum_{i=1}^n \sum_{j=1}^m w_{ij} h_j v_i - \sum_{j=1}^m c_j h_j - \sum_{i=1}^n b_i v_i \quad (1)$$

由已知可视层节点可以得到隐藏层的隐藏节点值,算式为

$$p(h_j = 1) = \frac{1}{1 + \exp(-b_j - \sum_i v_i w_{ij})} \quad (2)$$

同样地,由已知的隐藏层节点也可以得到可视层的节点值,即

$$p(v_i = 1) = \frac{1}{1 + \exp(-c_i - \sum_j h_j w_{ji})} \quad (3)$$

本质上,RBM网络是根据一个给定的输入信号  $\mathbf{v} = (v_1, v_2, \dots, v_n)$ , 根据网络输出对应的隐藏的特征向量  $\mathbf{h} = (h_1, h_2, \dots, h_m)$ , 使联合概率  $p(\mathbf{v}, \mathbf{h})$  最大。联合概率  $p(\mathbf{v}, \mathbf{h})$  满足

$$p(\mathbf{v}, \mathbf{h}) \propto \exp(-E(\mathbf{v}, \mathbf{h})) = e^{h^T \mathbf{W} \mathbf{v} + b^T \mathbf{v} + c^T \mathbf{h}} \quad (4)$$

## 1.2 RBM网络预训练

RBM网络预训练通过逐层训练多层RBM获得稳定的网络结构。可以通过最大化训练集的RBM网络的对数似然函数得到模型参数  $\theta$ , 即

$$\hat{\theta} = \operatorname{argmax}_{\theta} \sum_{i=1}^N \lg P(\mathbf{v}^{(i)} | \theta) \quad (5)$$

通过对每个训练样本进行Gibbs采样可以得到对应样本, 可以得到对数似然的梯度近似表达式

$$\frac{\partial \lg p(\mathbf{v} | \theta)}{\partial w_{ij}} = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}} \quad (6)$$

$$\frac{\partial \lg p(\mathbf{v} | \theta)}{\partial b_j} = \langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{model}} \quad (7)$$

$$\frac{\partial \lg p(\mathbf{v} | \theta)}{\partial c_i} = \langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{model}} \quad (8)$$

但是Gibbs采样次数较大, 尤其针对飞行高维数据时, 训练过程的效率十分低下。

2002年, HINTON提出对比散度(CD)快速学习方法。通常对比散度快速学习方法仅需要  $k$  步(一般仅需要1步), Gibbs提高计算速度以及保证计算的精度。参数更新如下, 即

$$\Delta w_{ij} = \varepsilon (\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}}) \quad (9)$$

$$\Delta a_i = \varepsilon (\langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{recon}}) \quad (10)$$

$$\Delta b_j = \varepsilon (\langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{recon}}) \quad (11)$$

式中:  $\varepsilon$  是学习速率;  $\langle \cdot \rangle_{\text{recon}}$  为样本分布的期望, 是Gibbs一步采样初始化数据得来的。

利用RBM自训练方法逐层从下到上无监督训练RBM结构。

## 1.3 微调阶段

BP神经网络是一个有监督的分类器, 在DBN中微调所有结构, 使整体达到最优。BP微调过程主要是利用误差逆传播训练的多层前馈网络, 不断地通过误差信号调整整个网络, 使网络达到整体最优。主要为以下两个步骤: 1) 信号从网络输入层逐层传播到网络输出层, 得到网络的输出信号。而其每个神经元都有一个激活函数, 一般地, 为Sigmoid非线性函数

$$y_i = \frac{1}{1 + \exp(-x_i)} \quad (12)$$

式中:  $x_i$  为神经元  $i$  的激活值,  $y_i$  为神经元  $i$  的输出值; 2) 通过网络输出值与标准值对比得到误差信号, 将信

号从输出层向输入层逐层传播, 优化DBN参数。

将BP网络与RBM结合, 起到微调整体DBN的作用, BP结合RBM网络训练步骤如下: 1) 随机初始顶层BP网络参数, 设定训练步长  $N$ ; 2) 信号从输入层向输出层逐层传播, 即前向计算, 对  $l$  层的  $j$  个神经元的值为  $y_j^l(n) = \sum w_{ij}^l(n) y_i^{l-1}(n)$ , 直到输出到输出层, 即  $l = L$ ,  $y_j^L(n) = O_j(n)$ , 误差信号  $e_j(n) = d_j(n) - O_j(n)$ ,  $d_j(n)$  为标准正确信号; 3) 通过反向传播计算局部梯度值, 对于输出层的梯度值为  $\delta_j^l(n) = e_j(n) O_j(n) [1 - O_j(n)]$ , 对于隐藏层的梯度值为  $\delta_j^l(n) = y_j^l(n) [1 - y_j^l(n)] \cdot \sum \delta_k^{l+1}(n) w_{kj}^{l+1}(n)$ ; 4) 更新权值  $w_{ij}^l(n+1) = w_{ij}^l(n) + \eta \delta_j^l y_i^{l-1}(n)$ ,  $\eta$  是学习速率; 5) 当  $n = N$  时, 结束; 否则  $n = n + 1$ , 转到步骤2)。

## 2 仿真试验

### 2.1 仿真试验数据与流程

试验数据采用某航空公司两组不同类别的QAR数据, 为了验证DBN特征提取技术的有效性, 将DBN与PCA网络提取的特征通过分类模型进行识别归类来评估本文提出的模型的有效性。仿真试验流程如下。

#### 2.1.1 数据预处理

第一组数据为同一航线同一机型在不同天气条件下的进近落地飞行数据, 一类为正常天气飞行数据, 即天气晴, 能见度大于10000m的飞行数据, 样本容量为  $300 \times 6160$ , 即共采集300组飞行数据, 每组为6160维飞行数据; 另一类为异常天气飞行数据, 即雨天, 能见度小于5000m的飞行数据, 样本容量为  $300 \times 6160$ 。

第二组数据为同一航线同一机型的不同飞行经验飞行员进近落地飞行数据, 一类为飞行小时在10000h以上的飞行员的飞行数据, 样本容量为  $300 \times 6160$ ; 另一类为飞行小时在5000h以上的飞行员的飞行数据, 样本容量为  $300 \times 6160$ 。

将飞行数据进行归一化处理。

#### 2.1.2 特征提取

分别采用PCA与DBN模型将飞行数据提取到2~10维飞行特征。

DBN模型的结构根据已有研究以及飞行数据特点, 本文选择了有5层DBN模型<sup>[9]</sup>进行仿真试验。节点数为6160-3000-1000-500-2~10, 即第1层是网络输入层为6160个节点, 因为飞行数据的维数为6160; 第2层为3000个节点; 第3层为1000个节点; 第4层为500个节点; 第5层为2~10个节点, 即DBN的输出层的节点为2~10个, 可以提取飞行数据2~10维特征。以正态随机分布初始化DBN模型, 其阈值初

始化为 0。RBM 最大迭代次数为 200,学习率为 0.15,动量参数为 0.95。

同样的,采用 PCA 算法将飞行数据提取到 2 ~ 10 维特征数据。

### 2.1.3 分类模型

用高斯过程分类器进行分类,核函数选用高斯核函数<sup>[10]</sup>。采用  $k$  折交叉模型对数据集进行训练, $k = 5$ ,将数据随机分为 5 组,进行 5 折交叉验证,使用 4 组训练高斯模型,剩下 1 组进行验证高斯模型分类结果的精度。反复进行 5 次,取分类准确度的平均值。

## 2.2 仿真结果与分析

为了验证模型正确性,利用 64 位 Windows10 操作系统,仿真软件 Matlab2015a 构建试验环境。首先将两组数据分别进行 PCA 与 DBN 特征提取,提取到 2 ~ 10 维,将 PCA 与 DBN 提到 3 维的特征进行可视化,图 3、图 4 所示为通过高斯分类器之前典型的特征数据。

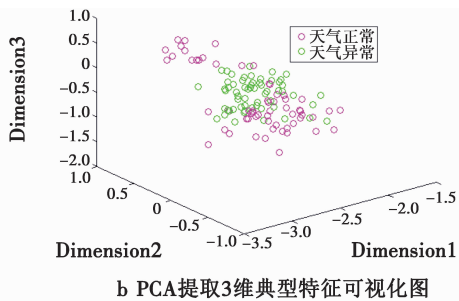
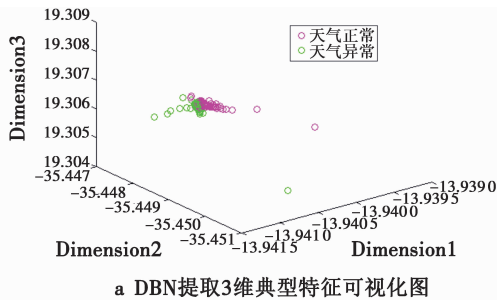


图 3 第 1 组数据的不同提取特征方法结果比较

Fig. 3 Feature extraction results of different methods on first set of data

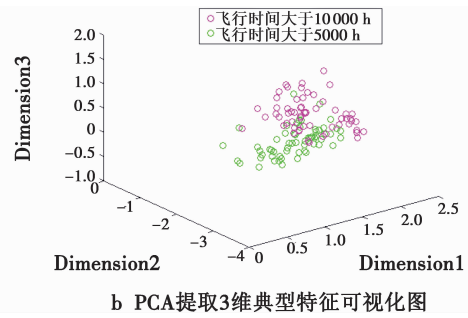
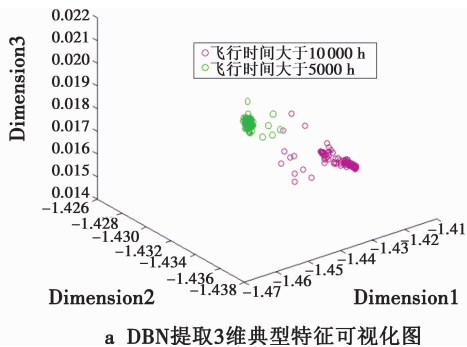


图 4 第 2 组数据的不同特征提取方法结果比较

Fig. 4 Feature extraction results of different methods on second set of data

从图中可以看到,PCA 提取第 1 组与第 2 组数据的特征不同类型的交叉重叠在一起的特征点较多,相互交错难以区分。从图 3a 与图 4a 可以看出,因为 DBN 最后一层是有监督的 BP 网络,可以很好地提取两类数据的差异性特征。通过 DBN 提取出来的特征,同一类别的特征有效地聚在一起,不同类别的特征有少量重叠,但是也可以很好地区分开来。把由 PCA 提取的特征图与由 DBN 提取的特征图对比可知,DBN 可以自适应地从原始飞行参数提取两类数据的差异性大的特征。而为了对比两类特征提取方法,通过表 1、表 2 分别提取 2 组数据到 2 ~ 9 维特征,且分别计算第 1 组数据第 1 类、第 2 类方差与第 2 组数据第 1 类、第 2 类方差,可以发现,采用 DBN 特征提取技术,同类数据方差明显小于采用 PCA 特征提取的同类数据方差。

表 1 第 1 组数据分别采用 PCA 与 DBN 特征提取的数据方差值

Table 1 Variance of the first set of data via PCA and DBN feature extraction methods

特征提取维数	PCA 第 1 类数据总方差	PCA 第 2 类数据总方差	DBN 第 1 类数据总方差	DBN 第 2 类数据总方差
2	313.37	112.62	0.00212	0.00201
3	315.24	154.72	0.00234	0.00213
4	382.95	160.05	0.00253	0.00243
5	392.34	195.17	0.00258	0.00246
6	420.12	200.27	0.00278	0.00263
7	439.88	204.11	0.00290	0.00279
8	460.34	204.92	0.00321	0.00301
9	464.24	219.90	0.00333	0.00314
10	481.80	220.01	0.00347	0.00378

为了对比两种方法的特征提取效果,本文将 PCA 特征提取后的数据与 DBN 特征提取后的数据进行分类,采用  $k$  折交叉模型对数据集进行训练,当  $k = 5$  时,将数据随机分为 5 组,进行 5 折交叉验证,使用 4 组训练高斯模型,剩下 1 组进行验证高斯模型分类结果的精度。反复进行 5 次,取分类准确度的平均值。通

过表3与图5可以得到,与通过PCA特征提取后的数据对比,不管是第1组数据还是第2组数据,通过采用DBN特征提取的数据分类准确率都高于PCA特征提取后的分类准确率。在第1组数据分类中,DBN特征提取到10维得到了0.912的分类正确率,而PCA提取10维特征只得到了0.811的分类准确率。在第2组数据分类中,DBN提取10维特征数据得到了0.922的分类正确率,而PCA提取到10维特征得到了0.791的分类正确率。验证了所提出的通过DBN对QAR飞行数据进行特征提取模型的效能。

表2 第2组数据分别采用PCA与DBN特征提取后的每类数据方差值

Table 2 Variance of the second set of data via PCA and DBN feature extraction methods

特征提取维数	PCA第1类数据总方差	PCA第2类数据总方差	DBN第1类数据总方差	DBN第2类数据总方差
2	273.4726	98.2819	0.00123	0.00134
3	283.3715	112.2718	0.00133	0.00141
4	289.3661	133.2819	0.00152	0.00169
5	327.3627	169.3618	0.00178	0.00177
6	353.2728	183.2719	0.00197	0.00180
7	361.2981	190.3761	0.00213	0.00232
8	379.9283	219.3718	0.00225	0.00241
9	381.3718	220.3718	0.00253	0.00248
10	391.3991	228.2761	0.00267	0.00281

表3 两组数据分别采用PCA与DBN特征提取后分类正确率

Table 3 Classification accuracy via PCA and DBN feature extraction

特征提取维数	第1组		第2组	
	PCA-GP分类准确率	DBN-GP分类准确率	PCA-GP分类准确率	DBN-GP分类准确率
2	0.634	0.734	0.672	0.712
3	0.676	0.756	0.689	0.721
4	0.692	0.786	0.723	0.801
5	0.712	0.823	0.732	0.791
6	0.721	0.845	0.719	0.821
7	0.745	0.876	0.743	0.834
8	0.789	0.861	0.773	0.854
9	0.765	0.901	0.782	0.878
10	0.811	0.912	0.791	0.922

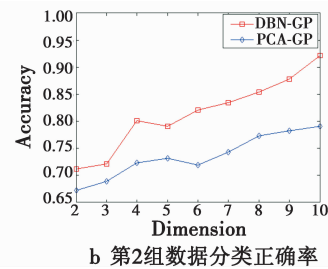
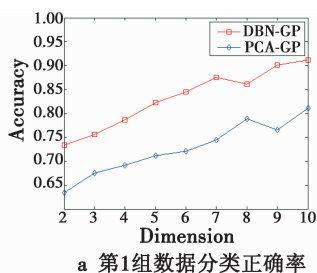


图5 通过PCA与DBN特征提取进行GP分类的准确度

Fig. 5 GP classification accuracy results via PCA and DBN feature extraction

### 3 结论

本文根据飞行数据大样本高维度的特点提出了一种有效的特征提取技术,该方法可以将飞行数据特征提取与分类识别技术结合在一起。在真实飞行数据集上进行仿真试验,通过DBN提取的特征进行分类,其分类准确度高于基于PCA特征提取技术的分类准确度,从而验证了该算法的有效性。

### 参考文献

- [1] 孙同江. 飞行数据的应用研究[D]. 南京:南京航空航天大学,2003.
- [2] NAIKAL N, YANG A Y, SASTRY S S. Informative feature selection for object recognition via sparse PCA[C]//IEEE International Conference on Computer Vision, 2012:818-825.
- [3] 邱宗江,刘慧霞,席庆彪,等. 无人机PCA故障检测与诊断技术研究[J]. 计算机工程与应用,2013,49(4):262-266.
- [4] LI L S, GABRIEL M, HANSMAN R J, et al. Anomaly detection in onboard-recorded flight data using cluster analysis[C]//IEEE/AIAA 30th Digital Avionics Systems Conference (DASC), 2011:213-226.
- [5] LECUN Y, BENGIO Y, HINTON G E. Deep learning[J]. Nature, 2015, 521(14539):436-444.
- [6] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786):504-506.
- [7] SALAKHUTDINOV R, MURRAY I. On the quantitative analysis of deep belief networks[C]//International Conference on Machine Learning, ACM, 2008:872-879.
- [8] ACKLEY D H, HINTON G E, SEJNOWSKI T J. A learning algorithm for Boltzmann machines[J]. Cognitive Science, 1985, 9(1):147-169.
- [9] BENGIO Y, LAMBLIN P, POPOVICI D. Greedy layer-wise training of deep networks[C]//NIPS06:Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems, Cambridge:MIT Press, 2007:53-160.
- [10] RASMUSSEN C E. Gaussian processes in machine learning[M]. Berlin:Springer Berlin Heidelberg, 2004.