

# 稀疏自编码深度神经网络及其在滚动轴承故障诊断中的应用

汤芳<sup>1</sup>，刘义伦<sup>1,2</sup>，龙慧<sup>1</sup>

(1.中南大学 机电工程学院，长沙 410083;2.中南大学 轻合金研究院，长沙 410083)

**摘要：**针对目前滚动轴承故障诊断主要采用监督式学习提取故障特征的现状，提出了一种基于稀疏自编码的深度神经网络，实现非监督学习自动提取滚动轴承振动信号的内在特征用于滚动轴承故障诊断。首先，将轴承故障振动信号的频谱训练稀疏自编码获得参数；然后用稀疏自编码获得的参数和轴承振动信号频谱的频谱训练深度神经网络，并结合反向传播算法对深度神经网络进行整体微调提高分类准确度；最后用训练好的深度神经网络来识别滚动轴承故障。对正常轴承、外圈点蚀故障、内圈点蚀故障和滚动体裂纹故障振动信号的分析结果表明：相比反向传播神经网络，提出的深度神经网络更能准确的识别滚动轴承故障类型。

**关键词：**稀疏自编码；深度神经网络；滚动轴承；故障诊断

**中图分类号：**TH133.3；TP206.3

**文献标识码：**A

## Deep Neural Network of Sparse Auto-encoder and Its Applications in Rolling Bearing Fault Diagnosis.

Tang Fang<sup>1</sup>, Liu Yilun<sup>1,2</sup>, Long Hui<sup>1</sup>

(1. College of Mechanical and Electrical Engineering, Central South University Changsha 410083, China

2. Light Alloy Research Institute, Central South University Changsha 410083, China)

**Abstract:** To overcome the problem of using supervised learning to extract fault features for classification in most of current rolling bearing fault diagnosis to extract the fault characteristics, a deep neural network algorithm is proposed, which is realized sparse auto-encoder, to achieve unsupervised feature learning by automatic extracting the inherent characteristics of the rolling bearing vibration signal for fault diagnosis of rolling bearing fault diagnosis. First, using spectrum of the bearing vibration signal to train sparse auto-encoder in order to obtain parameters; secondly, using parameters from sparse auto-encoder and spectrum of the rolling bearing vibration signal to train the deep neural network, the back-propagation algorithm is used for fine-tuning the deep neural network with the purpose of improving classification accuracy. Finally, the deep neural network, which has been trained to identify fault of rolling bearings. The analysis results from vibration signals with roller normal condition of the rolling bearing, pitting fault of bearing outer ring, pitting fault of bearing inner ring and crack fault of bearing rolling element show that, compared with back propagation neural network, the proposed deep neural network can identify fault type of rolling bearing fault accurately.

**Keywords:** sparse auto-encoder; deep neural network; rolling bearing; fault diagnosis

<sup>1</sup>滚动轴承在机械设备中应用广泛，也是极易出现损伤的部件之一，滚动轴承运行状况会影响整个机器的运行性能，对轴承进行状态监测和故障诊断具有重要的现实意义<sup>[1]</sup>。目前，滚动轴承的故障诊断主要是通

过局部特征尺度分解<sup>[2]</sup>、经验模态分解<sup>[3]</sup>、局部均值分解<sup>[4]</sup>等信号处理方法信号处理采集轴承的振动信号提取特征或将提取的特征作为样本训练反向神经网络（Back propagation neural network, BPNN）等神经网络<sup>[5-6]</sup>、支持向量机(Support vector machine, SVM)<sup>[7-8]</sup>等分类器完成轴承故障诊断。大多数基于机器学习算法<sup>[9-10]</sup>的诊断方法都是监督学习，这种算法需要大量标签数据和人为提取特征，标签数据作为“导师”监督训练分类器，人为特征提取需要研究者的经验和反复进行调整。

<sup>1</sup>收稿日期：2016-12-13

**基金项目：**国家自然科学基金(51375500, 61402167)和湖南科技大学机械设备健康维护湖南省重点实验室开放基金(201605)资助。

**作者简介：**汤芳 (1987-)，硕士研究生，研究方向为信号处理与机械故障诊断，[thstn@126.com](mailto:thstn@126.com);

**\*通讯作者：**刘义伦，教授，博士，[yiliu@csu.edu.cn](mailto:yiliu@csu.edu.cn)。

1986年, Rumelhart提出自动编码器<sup>[11]</sup>, 并将其用于高维复杂数据处理。2006年, Hinton提出了深度学习神经网络的概念及训练策略, 对原型自动编码器结构进行改进, 产生了深度自编码器<sup>[12]</sup>; 在此基础上, Bengio提出稀疏自动编码器的概念<sup>[13]</sup>, 深化了深度编码器的研究。稀疏自编码器能够利用无监督学习方式, 从无标签数据, 即从输入数据自身学习自动提取出数据特征, 减少人为提取特征的过程, 提高分类的精度, 稀疏自编码器和深度神经网络被广泛应用于图像分类<sup>[14-15]</sup>、语音识别<sup>[16-17]</sup>、故障诊断<sup>[18-19]</sup>等领域中, 取得了良好的识别效果。

本文利用稀疏自编码能自动提取学习特征的特性, 提出了一种基于稀疏自编码的深度神经网络, 实现非监督学习自动提取滚动轴承故障信号的学习特征用于滚动轴承故障诊断。用稀疏自编码获得权值和偏差初始化深度神经网络分类器, 将滚动轴承归一化后的频域数据作为样本训练和测试深度神经网络。通过分别对正常轴承、外圈点蚀故障、内圈点蚀故障、滚动体裂纹故障四振动信号进行分析表明, 提出的深度神经网络能有效的识别滚动轴承故障类型。

## 1 稀疏自编码基本原理

无监督学习<sup>[17]</sup>能够无标签数据中自主学习数据的内在特征, 简化了繁琐的人为提取特征步骤。稀疏自编码器作为一种典型的无监督学习算法, 可用于深度神经网络的建立并有效提取反映滚动轴承故障的内在特征。

自动编码器是一种利用无监督逐层贪婪与训练和优化系统参数的从无标签数据中提取高维复杂输入数据的分层特征, 并得到原始数据的分布特征表示的三层对称深度学习神经网络<sup>[20]</sup>, 结构如图1所示, 图中只显示了一个样本输入。自编码器由输入层、隐层和输出层组成, 隐层对原始输入数据进行编码, 输出层对隐层表达进行解码重构原始数据, 使得重构误差最小获得最佳的隐层表达。

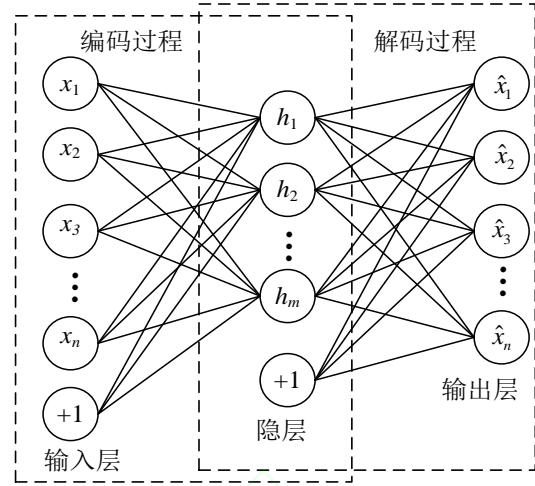


图1 自动编码器模型结构

对于采集到的滚动轴承故障状态振动数据  $X=\{x(1), x(2), x(3), \dots, x(k)\}$ , 其中  $x(i) \in \mathbf{R}^n$ ,  $k$  为无标签样本个数, 编码过程就是通过编码函数把每一个无标签训练样本为  $n$  维矢量  $x(i)$  映射成的  $m$  维编码矢量  $h(i)$ 。一般选择 sigmoid 函数作为编码函数, 然后编码矢量  $h(i)$  通过解码函数映射回重构矢量  $\hat{x}(i)$ , 解码函数一般为线性函数或 sigmoid 函数, 使得重构误差最小, 重构误差为

$$L(x(i), \hat{x}(i)) = \frac{1}{k} \sum_{i=1}^k \|x(i) - \hat{x}(i)\|^2 \quad (1)$$

如果编码矢量  $h(i)$  通过解码得到的重构矢量  $\hat{x}(i)$  与原始输入矢量的误差较小, 可以认为它保留了训练样本数据中大部分信息, 但是只保留原始数据  $x(i)$  的信息, 并不足以让自动编码器学习到有用的特征, 因为轴承所运行的工作环境复杂, 样本数据往往含有强烈的背景噪声, 因此需要给自编码器一定限制, 可以在自编码器隐层单元加入稀疏限制, 控制隐层单元激活数量, 隐层加入稀疏限制的自编码器称为稀疏自编码器。对于激活函数为 sigmoid 函数, 如果一个隐层单元输出值接近 1, 则这个单元处于“活跃”状态; 如果一个隐层单元的输出值接近 0, 则这个单元处于“不活跃”状态。我们希望隐层单元在大多数时候是处于“非活跃”状态。假设  $a_j(x)$  表示隐层第  $j$  个单元的激活量, 那么第  $j$  个单元的平均激活量可表示为

$$\hat{\rho}_j = \frac{1}{k} \sum_{i=1}^k a_j(\mathbf{x}(i)) \quad (2)$$

为了使大多数神经元为“非激活”状态，让  $\hat{\rho}_j$  等于一个接近 0 的常数  $\rho$ ， $\rho$  称为稀疏常数，在自编码器代价函数加入稀疏惩罚项，来惩罚  $\hat{\rho}_j$  偏离  $\rho$ 。选用KL散度<sup>[21]</sup>作为惩罚项PN的表达式

$$PN = \sum_{j=1}^{S_2} KL(\rho \parallel \hat{\rho}_j) \quad (3)$$

式中， $S_2$ 为隐层的单元数， $KL(\rho \parallel \hat{\rho}_j)$ 为KL散度。

KL散度数学表达式为

$$KL(\rho \parallel \hat{\rho}_j) = \rho \ln \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \ln \frac{1 - \rho}{1 - \hat{\rho}_j} \quad (4)$$

当  $\hat{\rho}_j = \rho$  时，KL 散度值为 0，KL 散度随着  $\hat{\rho}_j$  偏离  $\rho$  逐渐增大。对于自编码器而言，一般代价函数可以写为

$$J(\mathbf{W}, \mathbf{b}) = \frac{1}{k} \sum_{i=1}^k \|\mathbf{x}(i) - \hat{\mathbf{x}}(i)\|^2 + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{S_l} \sum_{j=1}^{S_{l+1}} (W_{ij}(l))^2 \quad (5)$$

式中， $\lambda$ 是权值衰减常数， $n_l$ 为神经网络层数， $S_l$ 为 $l$ 层的神经元个数， $S_{l+1}$ 为 $l+1$ 层神经元个数。则包含稀疏惩罚项的总代价函数可以写为

$$J_{sparse}(\mathbf{W}, \mathbf{b}) = J(\mathbf{W}, \mathbf{b}) + \beta PN \quad (6)$$

式中， $\beta$ 为稀疏惩罚项系数。

利用文献[12]提出的反向传播算法结合梯度下降法，更新连接权值  $\mathbf{W}$  和偏差  $\mathbf{b}$ 。求出使得式 (6) 达到最小时的  $\mathbf{W}$  和  $\mathbf{b}$ 。更新方程为

$$W_{ij}(l) = W_{ij}(l) - \alpha \frac{\partial}{\partial W_{ij}(l)} J_{sparse}(\mathbf{W}, \mathbf{b}) \quad (7)$$

$$b_i(l) = b_i(l) - \alpha \frac{\partial}{\partial b_i(l)} J_{sparse}(\mathbf{W}, \mathbf{b}) \quad (8)$$

式中， $\alpha$  为学习率。通过这种更新权值的方式，获得最优的  $\mathbf{W}$  和  $\mathbf{b}$ ，使得稀疏自编码隐层学习比较好的隐层表达。

## 2 基于稀疏自编码深度神经网络故障诊断方法

本文将无监督学习的稀疏自编码和监督学习的深度神经网络结合起来。利用稀疏自编码初始化深度神经网络，利用稀疏自编码隐层的特征表达训练深度神经网络，训练和微调整个深度神经网络。隐层获得的特征表达是连接权值  $\mathbf{W}$  和偏差  $\mathbf{b}$  的函数，因此可以把最优的  $\mathbf{W}$  和偏差  $\mathbf{b}$  作为深度神经网络的参数对深度神经网络进行初始化，获得有标签数据有效的特征表达。

文献[22]指出在深度神经网络获得好的性能方面，深度神经网络中的隐层单元数量比学习算法和自编码器深度更加重要。因此，本文只研究单隐层的深度神经网络，深度神经网络方法流程图如图 2 所示，步骤总结如下：

- 1) 采集正常状态、外圈故障、内圈故障和滚动体故障振动信号，获取振动信号的频谱，即频域数据。
- 2) 在原始数据样本随机选出一部分样本作为训练样本，剩余的作为测试样本。对训练样本和测试样本分别进行特征标准化预处理，使得频域数据每一个维度都具有零均值和单位方差。
- 3) 用预处理后的训练样本以无监督的方式训练稀疏自编码器，利用批量训练和梯度下降法获得  $J_{sparse}(\mathbf{W}, \mathbf{b})$  极小时的最优隐层参数  $\mathbf{W}$  和  $\mathbf{b}$ 。
- 4) 将训练稀疏自编码器的最优隐层参数  $\mathbf{W}$  和  $\mathbf{b}$  作为深度神经网络的隐层初始权值矩阵  $\mathbf{W}$  和偏差  $\mathbf{b}$ ，完成深度神经网络的初始化。
- 5) 根据轴承故障类型选用深度神经网络输出，本文选用 Softmax 分类器作为深度神经网络的输出。
- 6) 用训练样本以监督的方式，利用 BP 算法微调整个深度神经网络使  $J(\mathbf{W}, \mathbf{b})$  极小，完成整个深度神经网络的训练。
- 7) 将有标签的测试样本输入训练好的深度神经网络，利用深度神经网络进行轴承故障诊断和状态识别。

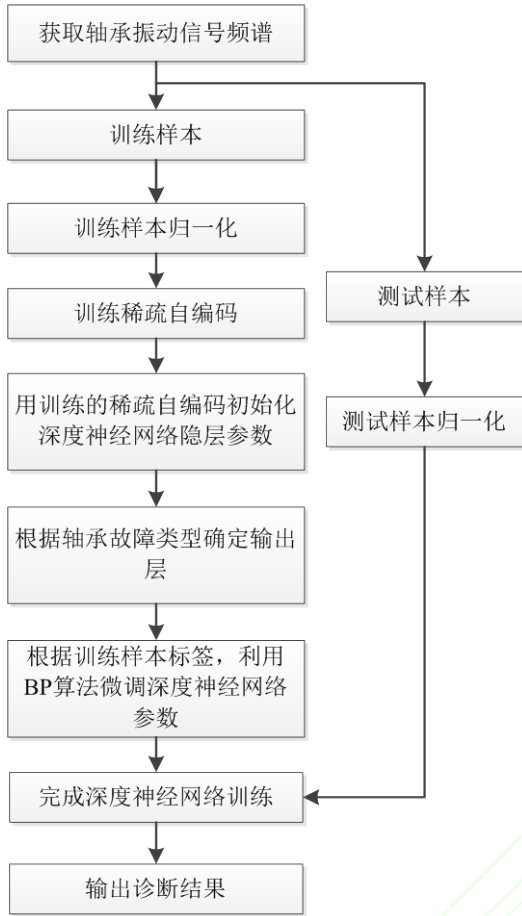


图2 提出方法流程图

### 3 实验研究

试验使用轴承型号为 N205EM, 轴的转速为 1200rpm, 采样频率为 10240Hz。振动信号由安装在轴承座上的加速度传感器来拾取。对正常轴承、外圈点蚀故障、内圈点蚀故障、滚动体裂纹故障四种轴承状态分别采样, 其中内、外圈点蚀故障是在外圈和内圈上加工一个直径为 0.5mm 的孔; 线切割技术分别在轴承内圈和外圈加工了一条宽为 0.5mm, 深为 0.5mm 的槽。采样点数是 102400 个数据点。将每种轴承状态采集的数据分为 100 组, 每种状态 1024 个数据点, 共 400 组数据。每种状态随机取一组数据, 得到如图 3 的振动波形图, 由图 3 可知, 当轴承发生故障时, 振动信号会表现出冲击特征。

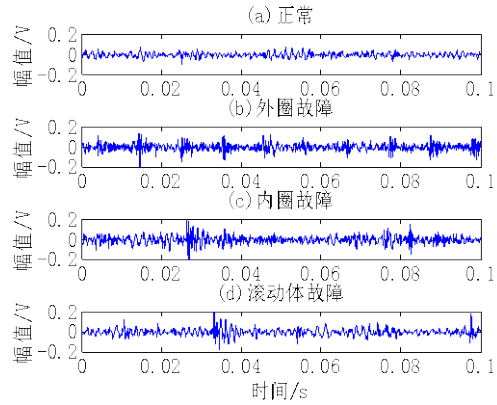


图3 轴承四种状态的波形图

将图 3 中的数据进行傅立叶变换得到轴承不同状态下的频谱图, 如图 4 所示, 为便于观察轴承发生故障时的频谱, 只显示了频率范围[0 2000]的频谱。由图 4 可知, 轴承发生故障时, 振动信号的频谱会表现出与故障部位相对应的故障频率及其倍频。从图 3 和图 4 可以看出, 滚动轴承运行环境十分复杂, 采集到滚动轴承振动信号含有大量的强烈背景噪声, 振动数据中的背景噪声掩盖了振动数据有用的信息—周期冲击特征和故障特征频率。

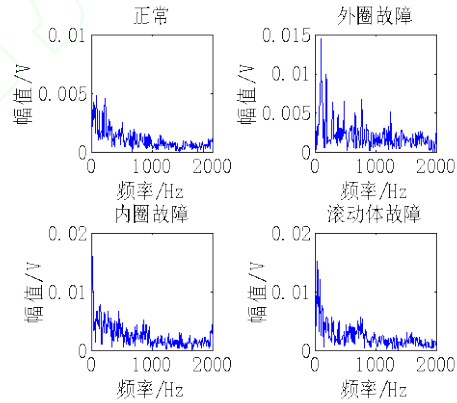


图4 轴承四种状态的频谱图

本文利用稀疏自编码来自动提取轴承故障数据的内在特征并完成神经网络的训练, 对轴承故障类型进行识别。将上述 400 组数据的频谱, 即频域数据, 将轴承不同状态下的频域数据作为样本。实验中, 随机选出 300 组数据作为训练样本, 剩余的 100 组数据作为测试样本。稀疏自编码的结构为: 513-256-513; 深度神经网络的结构为: 513-256-4。

为验证提出方法有效性, 用与深度神经网络结构、学习算法和激活函数等均相同的单隐层 BPNN 和



双隐层 BPNN 诊断相同的数据集，为了减少随机因素影响，试验重复进行共 15 次，每次试验迭代次数为 50 次。图中，BP256 代表单隐层 BPNN，隐层神经单元数为 256。BP256-128 表示双隐层 BPNN，隐层神经单元数分别为 256 和 128。试验结果如图 4 所示，从图中可以看出，15 次试验的诊断准确率都为 100%，这说明提出的方法能排除背景的干扰，准确的识别轴承故障类型。对于双隐层 BPNN 和单隐层 BPNN 诊断准确率在 20%~60%，诊断准确率波动较大且准确率整体较低。

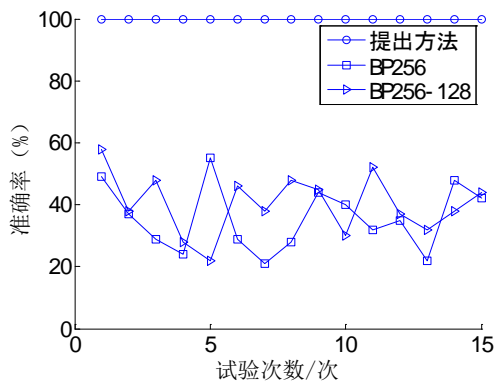


图 5 15 次试验诊断准确率

从图 5 可以看出，第 7 次试验的与深度神经网络结构相同的单隐层 BPNN 诊断率最低。以第 7 次试验中的结果进行具体分析，图 6 为第 7 次三种训练方法训练误差，从图中可以看出。单层 BPNN 训练方法在第 11 次迭代时陷入了局部最优，导致训练误差保持在 0.2504 左右，诊断准确率为 21%；而双隐层 BPNN 训练方法训练误差随着迭代次数逐渐降低，尚未陷入局部最优。但训练误差下降速率缓慢，到迭代结束时训练误差为 0.2409，低于单层 BPNN 训练误差，仍在较高的水平。因此诊断准确率为 38%，高于单层 BPNN 方法准确率，但是准确率仍不高。深度神经网络微调是在无监督学习方式的稀疏自编码预训练基础上进行的，训练开始时训练误差就很小，第 1 次迭代训练误差为 0.0005，已接近 0。误差下降速率较快最后 1 次迭代时训练误差下降至 0.00004，稀疏自编码的预训练深度神经网络有效避免深度神经网络陷入局部最优。因此，相对于 BP 神经网络，稀疏自编码深度神经网络具有更好的诊断能力。

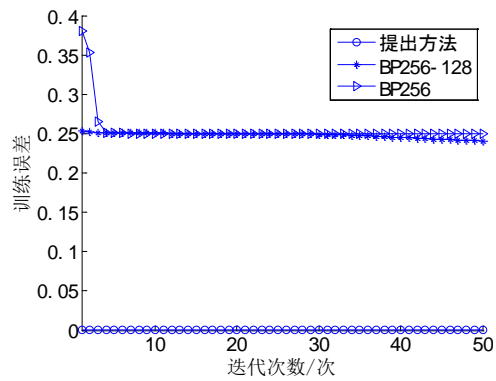


图 6 三种方法训练误差

为更好比较三种方法的精度，计算 15 次随机试验的诊断准确率的平均值及相应的标准差，同时研究不同分类器对稀疏自编码自动学习的隐层特征的分类能力，试验中，将深度神经网络与 SVM 和 Softmax 的分类能力进行比较，这两种分类器只对稀疏自编码器学习的隐层表达进行分类，不对整个网络进行微调，结果表 1 所示。从可以看出，基于稀疏自编码的深度神经网络能有效进行轴承故障诊断，诊断准确率大大高于 BPNN 方法；无论使用浅层的 BPNN 方法还多层 BPNN 方法，轴承故障诊断精度都不高且标准差在 10%左右。综合 Softmax、SVM 以及本文提出方法诊断正确率，最低准确率已超过 80%，标准差最大为 5.05%，由此可以说明进行无监督学习的稀疏自编码能有效的发现轴承故障振动频域数据的内在特征。对比后三行可以发现，深度神经网络可以稀疏自编码提取的特征进行调整，其诊断精度要高于在稀疏自编码器基础上不进行微调 Softmax 分类器和 SVM 分类器。对于轴承故障诊断，深度神经网络具有更好的诊断性能。

表 1 15 次试验的诊断结果

方法	平均准确率/%	准确率标准差/%
BP256	35.67	10.47
BP256-128	40.27	9.73
Softmax	88.53	4.71
SVM	95.8	5.05
提出方法	100	0

## 4 结论

通过稀疏自编码初始化的深度神经网络可以有效进行轴承故障诊断，获得数据内在特征。通过对比研究可以得到以下结论：

1)稀疏自编码能够以无监督学习的方式获取原始数据的内在特征,即在无标签条件下,自动获得原始数据的内在有效特征。

2)深度神经网络能对稀疏自编码学习的特征进行微调,与 SVM 以及 Softmax 相比,深度神经网络能有效地诊断轴承故障。

3)提出的基于稀疏自编码深度神经网络,与 BPNN, 有更强的优化参数的能力,降低训练误差,更适合进行轴承故障诊断。

## 参考文献

- [1] 王国彪,何正嘉,陈雪峰,等. 机械故障诊断基础研究“何去何从”[J]. 机械工程学报, 2013, 49(1): 63-72.  
WANG G B, HE Z J, CHEN X F *et al.* Basic research on machinery fault diagnosis-what is the prescription [J]. Journal of Mechanical Engineering, 2013, 49(1): 63-72 (in Chinese)
- [2] Liu H, Wang X, Lu C. Rolling bearing fault diagnosis based on LCD-TEO and multifractal detrended fluctuation analysis[J]. Mechanical Systems and Signal Processing, 2015, 60: 273-288.
- [3] Zheng J. Rolling bearing fault diagnosis based on partially ensemble empirical mode decomposition and variable predictive model-based class discrimination[J]. Archives of Civil and Mechanical Engineering, 2016, 6(4): 784-794.
- [4] 陈鹏飞,陈卫,高星伟,等. LMD 和支持向量机相结合的齿轮毂故障诊断方法[J]. 机械科学与技术, 2015, 34(10): 1599-1603.  
Chen P F, Chen W, Gao X W, *et al.* Fault diagnosis of gear hub based on LMD and support vector machine[J]. Mechanical Science and Technology for Aerospace Engineering, 2015, 34(10): 1599-1603 (in Chinese)
- [5] 姚德臣,贾利民,秦勇,等. 基于小波 BP 神经网络的城轨列车辅助逆变系统故障诊断[J]. 中国工程机械学报, 2013, 11 (6): 542-546.  
YAO D C, JIA L M, QIN Y, *et al.* Fault diagnosis on auxiliary inverter system of urban rail vehicles based on wavelet BP neural network[J]. Chinese Journal of Construction Machinery, 2013, 11 (6): 542-546 (in Chinese)
- [6] 杨一舟,蒋东翔. 概率神经网络用于机匣振动故障诊断[J]. 机械科学与技术, 2016, 35(12): 1805-1810.  
Yang Y Z, Jiang D X. Casing vibration fault diagnosis based on probabilistic neural networks[J]. Mechanical Science and Technology for Aerospace Engineering, 2016, 35(12): 1805-1810. (in Chinese)
- [7] 李永波,徐敏强,赵海洋,等. 基于层次模糊熵和改进支持向量机的轴承诊断方法研究[J]. 振动工程学报, 2016, 29(1): 184-192.  
LI Y B, XU M Q, ZHAO H Y, *et al.* A study on rolling bearing fault diagnosis method based on hierarchical fuzzy entropy and ISVM-BT[J]. Journal of Vibration Engineering, 2016, 29(1): 184-192 (in Chinese)
- [8] 姜久亮,刘文艺,侯玉洁,等. 基于内积延拓 LMD 及 SVM 的轴承故障诊断方法研究[J]. 振动与冲击, 2016, 35(6): 104-108.  
JIANG J L, LIU W Y, HOU Y J, *et al.* Bearing fault diagnosis based on integral waveform extension LMD and SVM[J]. Journal of Vibration and Shock, 2016, 35(6): 104-108 (in Chinese)
- [9] 蔺想红,王向文,张宁,等. 脉冲神经网络的监督学习算法研究综述[J]. 电子学报, 2015, 43(3): 577-586.  
LIN X H, WANG X W, ZHANG N, *et al.* Supervised learning algorithms for spiking neural networks: A review[J]. Acta Electronica Sinica, 2015, 43(3): 577-586 (in Chinese)
- [10] 石灵芝,邓启红,路婵,等. 基于 BP 人工神经网络的大气颗粒物 PM<sub>10</sub>质量浓度预测[J]. 中南大学学报(自然科学版), 2012, 43(5): 1969-1974.  
SHI L Z, DENG Q H, LU C, *et al.* Prediction of PM<sub>10</sub>-mass concentrations based on BP artificial neural network[J]. Journal of Central South University(Science and Technology), 2012, 43(5): 1969-1974 (in Chinese)
- [11] Williams D, Hinton G E. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.

- [12] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [13] Bengio Y, Lamblin P, Popovici D, et al. Greedy layer-wise training of deep networks[J]. Advances in neural information processing systems, 2007, 19: 153-160.
- [14] 钱彬, 唐振民, 徐威. 基于稀疏自编码的路面裂缝检测[J]. 北京理工大学学报, 2015, 35(8): 800-804.  
QIAN B, TANG Z M, XU Wei. Pavement crack detection based on sparse auto-encoder[J]. Transactions of Beijing Institute of Technology, 2015, 35(8): 800-804 (in Chinese)
- [15] Yin H, Jiao X, Chai Y, et al. Scene classification based on single-layer SAE and SVM[J]. Expert Systems with Applications, 2015, 42(7): 3368-3380.
- [16] 马勇, 鲍长春. 基于稀疏神经网络的说话人分割[J]. 北京工业大学学报, 2015, 41(5): 662-667.  
MA Y, Bao C C. Speaker segmentation based on sparse neural network[J]. Journal of Beijing University of Technology, 2015, 41(5): 662-667 (in Chinese)
- [17] Cheriadat A M. Unsupervised feature learning for aerial scene classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(1): 439-451.
- [18] Bengio Y. Learning deep architectures for AI[J]. Foundations and Trends in Machine Learning, 2009, 2(1): 1-127.
- [19] 孙文珺, 邵思羽, 严如强. 基于稀疏自动编码深度神经网络的感应电动机故障诊断[J]. 机械工程学报, 2016, 52(9): 65-71.  
SUN W J, SHAG S Y, YAN R Q. Induction motor fault diagnosis based on deep neural network of sparse auto-encoder[J]. Journal of Mechanical Engineering, 2016, 52(9): 65-71 (in Chinese)
- [20] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21): 49-56.  
LEI Y G, JIA F, ZHOU X, et al. A deep learning-based method for machinery health monitoring with big data[J]. Journal of Mechanical Engineering, 2015, 51(21): 49-56 (in Chinese)
- [21] Kullback S, Leibler R A. On information and sufficiency[J]. The Annals of Mathematical Statistics, 1951, 22(1): 79-86.
- [22] Coates A, Lee H, Ng A Y. An analysis of single-layer networks in unsupervised feature learning[J]. Ann Arbor, 2011, 15: 215-233.