Geometric Features Fusion Method based deep Learning for fault Classification

Funa Zhou1, Po Hu1,\*，Chenglin Wen2,\*

1.School of Computer and Information Engineering, Henan University, Kaifeng, China

2. School of Automation, Hangzhou Dianzi University, Hangzhou, China

[hupo4210@163.com, wencl@hdu.edu.cn (\* Corresponding](mailto:%20gaoyulinhn@163.com,%20wencl@hdu.edu.cn%20(*%20Corresponding) Author)

**Abstract**: 机械设备运行期间需要实时监测各种类型的故障，有些频率类故障无法在时域中被很好的检测出来，这就使得无法对机械设备做到精准的实时检测。针对这个问题，提出了一种基于DNN的几何特征融合的故障诊断方法。通过引入原始数据的几何特征解决了当频率类故障存在的幅值相近且频繁过零点带来时域检测不准确的问题。将原始、斜率、曲率数据分别构建自动编码器模型，充分挖掘数据隐含的频率信息。然后将得到的数据特征进行特征融合，其特征维度增大使得分类模型训练的更准确。滚动轴承实验结果表明本文提出的方法的能够很好的检测出频率类故障。

**Key words**: 深度神经网络 ；自动编码器；特征融合；故障诊断

**1. Introduction**

滚动轴承是旋转机械设备的核心部件，在大型的工业过程控制中应用极为广泛。由于轴承在使用过程中会出现不同程度的磨损现象，如果不能及时的监测到轴承发生的故障将会导致严重的损失，甚至危害到操作人员的生命[1-4]。因此在实际应用中对于轴承进行实时的故障诊断是非常有必要的。

目前机械设备故障诊断的方法主要有基于定性模型的方法、基于定量模型的方法和基于数据驱动的方法[5-7]。基于定性模型的和定量模型的方法需要对具体的诊断对象有深刻的了解并且能够抽象出准确的数学模型，所以这些方法太过于依赖专家的先验知识，以至于这些方法的泛化能力太差，在实际应用较为少见。基于数据驱动的方法有统计特征提取的方法和深度学习的方法[8-11]。基于统计特征提取的方法只能实现故障的检测，不能对故障进行分类。近些年来由于深度学习方法的兴起，基于深度学习的方法在复杂系统的故障诊断中被广泛应用，例如卷积神经网络、递归神经网络和深度神经网络等方法。

在轴承故障诊断中，振动信号是最容易采集到的，并且振动信号对轴承的故障比较敏感，因此振动信号被广泛用于轴承故障诊断中。但是，由于机械设备故障特征具有非线性、不稳定性、高纬度以及含有大量的噪音污染等特性，使得无法对机械设备进行精确的故障诊断[12-14]。有的学者提出了一些利用信号处理的特征提取方法技术与机器学习方法相结合的方式对机械设备进行故障分类。 Widodo and Yang提取频域特征作为SVM的数据源来检测机械故障[13]。 当样品的数量很少并且信号是非平稳时， Yu等提出了一种利用SVM和EMD方法相结合的方式进行滚动轴承故障分类[10]。Hu等从振动信号中提取小波包变换（Wavelet Packet Transform,WPT）各个节点能量作为轴承故障诊断的特征参数，极大地保留了特征信息的时频特性，利用WPT和SVM相结合的方式对轴承进行故障诊断，提高了故障诊断的精确度[15]。Wang等根据根据滚动轴承振动信号的非平稳特征，利用WPT对采集信号进行降噪，提取各频带小波棒的能量特征作为人工神经网络（Artificial Neural Network，ANN）的输入特征，该方法利用ANN非线性学习分类能力和自组织能力对轴承故障进行分类和诊断[16]。Yang和Tang等人提出一种利用专家系统和BP神经网络(Back Propagation Neural Network，BPNN)相结合的方法，该方法充分利用了专家系统和ANN的优势，成功地检测出轴承发生的故障[17]；Jiang等提出一种使用高阶累积量和BPNN相结合的方法，该方法利用高阶统计量作为特征向量，提高了BPNN对轴承进行故障诊断的准确率[18]；但是在这些研究中， SVM和BPNN作为浅层的学习方法存在着很多缺陷：SVM本质是一个二分类器，在多种分类及大样本问题上学习效率低，如何选择合适的核函数以及尺度参数等往往需要通过经验，SVM方法不能进行实时监控和诊断，无法满足当前对机械进行实时监测并进行故障诊断的要求；ANN也有一些先天性的缺陷：（1）ANN作为一种浅层神经网络，收敛速度比较慢，极易陷入局部最优，不能很好地表征信号的特征信息；（2）学习复杂非线性的数据存在着效率不高，分类精确度较低的缺点。综上所述，SVM和BPNN作为浅层学习方法，已经不能适应高维非稳态数据条件下对特征进行有效提取[19]。

深度学习是一个有前途的特征提取工具并且引起了学者们的广泛关注[21,28-30]。与浅层学习相比，深度学习可以很好地进行特征提取以及对非线性大数据问题进行深入研究[31,32]。深度神经网络（DNN）通过无监督的逐层贪心训练算法和基于BP的全局参数微调，不仅可以避免局部最优问题，也解决标签和样本数量的限制，提高泛化能力。 2006年Hinton和Salakhutdinov第一次提出了深度学习理论。该理论用于构建深度神经网络(Deep neural network,DNN)，通过多层的非线性变换，组合低层特征形成更加抽象的高层表示，使得一个学习系统能够不依赖人工的特征选择，发现数据的分布式特征表示，并学习到复杂的表达函数[11]。鉴于其优秀的特征提取能力，它很快引起了故障分类领域专家的重视。Lu等利用深度神经网络良好的特征提取能力，将其用于轴承故障诊断，成功解决了传统特征提取方法不能及时有效地发现未知类型故障的问题[33]；Jia等用深度神经网络来检测滚轴轴承的健康状态[34]；Gan等根据轴承故障发生的特点提出了一种分层深度神经网络的故障诊断方法，该方法通过构建一个两层的深度神经网络，不仅能够准确定位轴承发生故障的位置，而且能够有效的挖掘轴承在同一位置下的故障尺寸[11]。深度学习作为当今最流行、最热的机器学习方法，给人工智能领域带来了颠覆式的革命，但是我们对深度学习的应用还处于刚起步阶段，在应用的过程中还有很多可以改进的方面。比如深度学习方法对机械设备频率类故障是很难做到实时性的，虽然频率类故障在频域能够被很好的检测出来，但是无法保证检测的实时性，而在时域中检测频率类故障是非常困难的。因为要保证实时性，实际中就需要对设备进行实时的检测，因此在时域中如何有效的检测出频率类故障是十分必要的。

为解决上述问题，本文提出了一种基于DNN的几何特征融合的故障在线诊断方法。其主要创新点是通过引入原始数据的几何特征解决了当频率类故障存在的幅值相近且频繁过零点带来时域检测不准确的问题。将原始、斜率、曲率数据分别构建自动编码器模型，这样可以充分挖掘数据隐含的频率信息。然后将得到的数据特征进行特征融合，其特征维度增大使得分类模型训练的更准确。首先计算出原始数据的斜率、曲率等能表示数据频率特征的值。第二步是构建三个DNN网络用于每个特征的训练，第三步是将得到的特征进行融合，最后一步是利用已经训练好的网络对新来的样本进行故障的实时诊断。

文中剩下的部分组织结构如下：第二部分概述深度学习理论背景；第三部分提出一种基于DNN的几何特征融合的故障在线诊断方法；第四部分通过实验和仿真分析得到提出的故障诊断方法的有效性；最后是本文的结论部分。

**2 深度学习理论**

深度学习是一种无监督特征学习方法，通过多层的非线性变换，组合低层特征形成更加抽象的高层表示，使得学习系统可以不依赖人工的特征选择，发现数据的分布式特征表示，并学习到复杂的表达函数。深度学习先以非监督学习方式对 DNN 进行逐层预训练，帮助 DNN 有效挖掘机械信号中的故障特征；然后以监督学习方式对 DNN 进行微调，优化 DNN 对故障特征的表达并使其具有监测诊断能力。本文通过堆叠多个自动编码器（Autoencoder, AE）对DNN进行预训练。

2.1自动编码器

自动编码器是一种三层的无监督神经网络，分为编码网络和解码网络两个部分，如图1所示，主要有输入层、隐含层和输出层组成。AE 的输入数据和输出目标相同，通过编码网络将高维空间的输入数据转换为低维空间的编码矢量，通过解码网络将低维空间的编码矢量重构回原来的输入数据。由于在输出层可对输入

信号进行重构，使得编码矢量成为输入数据的一种特征表示。

****

Fig. 1 AE model structure

给定一个无标签数据集，包含个观测特征或样本，每个观测变量含有样本，编码网络通过激活函数将样本编码为隐含层的激活值，编码过程如式1所示：

 （1）

其中：为编码函数，函数通常取Sigmoid函数作为编码网络的激活函数，*W*为输入层和隐含层之间的网络权值矩阵，*b*为编码网络产生的偏置向量，为输入层和隐含层之间的连接权值和偏置参数，Sigmoid函数的一般形式如式2所示：

 （2）

同样地，对于解码网络，就是将编码网络得到的编码矢量通过解码网络重构得到和输入相等，即和输入相等，解码过程如图式3所示：

 （3）

式中，为解码函数，为编码过程的激活函数，为隐含层到输出层之间的网络权值矩阵，*d*为编码过程产生的偏置向量。

训练AE的过程实质就是对网络参数的训练和优化。为了使输出尽可能地和输入接近，我们需要对网络训练参数进行优化，通过最小化重构误差来刻画输入和输出之间的接近程度，如式4所示：

 （4）

在每一次训练的过程中，利用梯度下降法来更新AE的网络训练参数，整个参数更新过程如下

 （5）

 （6）

其中，为学习率， 和是利用反向传播算法计算出来的。

DNN可以简单看做是多个AE进行堆叠而成的多隐层神经网络，使用自下而上的无监督学习方法，对特征进行逐层提取，利用有监督的学习方法对整个网络参数进行微调，从而使DNN能够从原始数据数据中提取到物体某种状态的最本质特征属性，DNN结构如图2所示。

图2 DNN预训练过程

首先，通过无监督逐层贪心训练算法对DNN网络进行预训练。首先通过给定一个无标签输入数据集作为编码网络的输入来练第一个自动编码器AE1， 通过AE1得到编码矢量。训练参数通过设置作为AE1的输出得到。然后将作为第二个自动编码器AE2的输入并训练AE2的网络参数，作为AE2的隐含层数据可以被看作AE2的一个特征表示。重复这个过程，得到第个自动编码器AE *N*的隐含层特征和相应的网络训练参数。

其次，对DNN网络顶端添加分类器。通过逐层无监督训练的方法完成DNN的预训练过程，实现了对特征信息的层层提取。但是此时的DNN并不具有分类的功能，为了实现输出分类的功能，还需要在DNN的顶层添加一个分类器。本文使用Softmax分类器作为DNN的输出层，假设训练数据集是 , 标签是, 每种类型的概率可以通过下面的假设函数来计算：

 （7）

其中是Softmax 的模型参数，与AE模型类似，为了保证分类器的性能，通过最小化cost函数来优化模型的训练参数. Softmax训练过程cost函数如式8所示, 顶层网络参数可以通过最小化来得到。

.

 （8）

最后，DNN训练参数微调。为了特征提取的准确率和输出层的分类效果，通过有限个样本标签，利用反向传播算法对整个DNN训练参数进行有监督的微调，通过最小化重构误差来完成微调过程，参数更新的过程如下：

 （9）

 （10）

其中，是真是的输出值，是从整个网络训练的过程中得到的参数集合，反向传播算法用来更新网络参数，是学习率，微调过程利用带标签的数据提高DNN训练的精确性。

**2.2 DNN分类**

为了能够准确地利用DNN模型从输入样本中提取到机械设备健康状况的本质特征，需要以下几个步骤：首先，应该对采集到的振动信号进行数据预处理，由于本文主要针对频率类故障在时域中难以有效检测的问题，因此求出原始时域振动信号相应的斜率和曲率值，用以反应原始信号的频率信息。其次，使用预处理的数据作为DNN模型输入无监督层层预训练来提取机械设备健康状况的时频特征。最后，根据样本的有限个标签，利用反向传播算法对整个网络进行微调，更新整个网络参数。这样我们可以对机械装备健康状况进行一个有效的分类和诊断。机械装备的故障诊断通常分为两个过程：训练过程和测试过程。将预处理的数据集分为训练数据和测试数据，训练数据用于构建并训练DNN模型，得到训练参数，利用训练参数初始化测试数据，验证构建模型的诊断效果，将错误分类的个数作为DNN分类精确度的一个参考指标。DNN用于机械系统故障诊断的详细步骤如图3所示。



**图3.** 基于DNN的故障分类框图

**3．基于深度学习的频率类故障分类模型**

实际系统中有许多频率类故障。对于此类故障的处理，多数是对原始数据进行傅里叶变换得到频域信息后再处理，这样会导致无法做到实时检测，本文通过求出原始数据的斜率和曲率数据，来增加时域信号的频率信息。分别构建对应的自编码器来分别提取原始数据、斜率数据、曲率数据的特征，然后将特征进行融合并最终实现分类。该网络结构如图4所示。



Figure 4 network structure

频率类故障监测的详细步骤如下。

步骤1 计算原始数据的斜率和曲率数据

在这一步中，我们计算出每一个时刻原始数据相对应的斜率曲率数据。用以在时域中得到频率类信息。

步骤2 建立原始数据的DNN网络

我们建立一个DNN模型来分别提取原始数据的故障特征。特征提取过程详细说明如下：

(1) 构建一个具有N个隐层DNN网络DNN1，并初始化DNN1训练参数。

 (11)

其中，**是权值矩阵，**偏置向量. 是 *DNN*1隐层神经元的个数。 网络的配置保存在中。表示原始数据的训练数据集。使用式（12）中的来表示*DNN*1输入神经元的个数。

 (12)

通过式(13)-(14)初始化*DNN*1的参数

 (13)

 (14)

（2）通过训练*DNN*1获取参数，对数据集应用无监督逐层特征提取的方法

 (15)

 (16)

在*DNN*1顶层增加一个softmax分类器，有限个标签集被用于*DNN*1的反向调整和训练参数的更新，通过式(17)-(18).

 (17)

 (18)

其中, 可以通过(7)-(8)被计算出来，是样本数目 。 表示 *DNN*1的输出，是反向微调过程的学习率。

（3）利用训练好的*DNN*1进行故障的分类，用测试样本，通过训练好的网络可以计算出每个测试样本的概率，然后利用式（19）将输出测试样本的类别。

 （19）

其中;是样本的模态类型. **表示第m个测试样本的故障类别。

通过真实的故障标签和网络测试得到的故障分类标签计算错误分类的误分个数：

 (20)

其中，是特征数据集的大小，是错误分类数据集。

 (21)

步骤3 建立斜率数据的DNN网络

(1) 为了得到原始数据的频域信息，我们构建了第二个分层模型，它的训练数据是斜率数据，表示*DNN*2训练集。

 (22)

*DNN*2参数的初始化机制和步骤1相同。

(2) 训练*DNN*2获取网络参数，相似地，详细的计算过程参考式(15)-(18)。

(3) 通过训练得到的**的网络参数

 (23)

基于训练得到的**可以通过式（24）来进行斜率数据的故障诊断。

 (24)

通过与实际标签进行对比计算错误分类的个数

步骤4 建立曲率数据的DNN网络

为了进一步提取数据的斜率特征，设计了第三个DNN网络。构建第三个深层网络，是*DNN*3的训练数据集，是测试数据集。参数的训练处理类似个步骤2，在中第m个样本的故障类别可以通过式（26）计算出来。

 (26)

步骤4 特征融合。

在本文中，以原始、斜率、曲率数据分别建立的DNN的模型，可以充分提取故障信号中含有的频率类故障成分。为使得最终的诊断精度较高，需要充分利用已训练好的网络提取数据的频域和时域特征。特征融合是将这三个自动编码器抽取的高层特征进行重新组合。然后对重新组合后的特征利用训练好的softmax分类器进行分类。特征的组合方式如下：



Fig.2 New feature vector

本文提出的基于分层的深度神经网络的多模故障分类方法流程图如图5所示

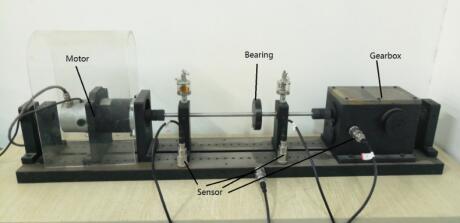
图5 基于几何特征的DNN的故障诊断流程图

1. **Experiment and Analysis of Bearing Faults**

Rolling bearings play a crucial role in rotating machinery, bearing health will directly affect the reliability and stability of the entire system. In this paper, the rolling bearing is used as an experimental platform to verify the effectiveness of the fault diagnosis method for geometric feature fusion. The proposed method is compared with the neural network method without feature fusion.

**4.1 Description of Experimental Platform**

The experimental data set used in this paper is collected by Henan University of Technology。Acceleration sensor is used in the experiment to collect the vibration signal of motor drive end as experimental data of bearing fault diagnosis. In this experiment, the acceleration sensor is used to collect the motor-driven vibration signal of 0hp, the sampling frequency is 48kHz. There are four types of fault diameter :(1)0.007、(2)0.014、(3)0.021 (4) normal condition. Fourier transform is used to preprocess the vibrations signal to get 4500 training samples and 4500 test samples.The experimental platform shown in Fig.4.



**Fig.2** **Rolling Bearing Experimental Platform**

**4.2 Data Description**

This paper is mainly to solve the problem of frequency-type fault detection in the time domain is difficult to effectively, when the assignment of similar fault signal to the depth of the traditional learning method is difficult to effectively identify different fault types. Therefore, this paper selects three sets of fault data and a set of normal data, the three sets of fault data are the inner fault type, and the load is 0, the difference is the fault diameter is 0.007,0.014,0.021, the data type shown in Table 1. The four data types are selected because these data types have similar amplitudes and frequent zero crossings, so that it is difficult to effectively distinguish the types of faults based on the time-domain signals only while ensuring that other operating conditions are consistent. To ensure real-time and can not be transformed in the frequency domain, this paper proposes to calculate its slope curvature data to increase its frequency domain features. Time domain signal shown in Fig.4 **.**



Fig 4 Observation of original signals

**Tab.1.** Four signal types

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mode | Load (hp) | Rotating speed(rpm) | Sampling frequency (S/s) | Fault diameter(mil) |
| Mode1 | 0 | 500 | 12000 | 0.007 |
| Mode2 | 0 | 500 | 12000 | 0,014 |
| Mode3 | 0 | 500 | 12000 | 0,021 |
| Mode4 | 0 | 500 | 12000 | Normal |

**4.3 Analysis of results**

Our proposed deep learning based on geometric features is applied to bearing fault diagnosis. There are 4,500 samples under each data type and 4 different data types to characterize the frequency fault types of rotating mechanical systems. In order to reduce the influence of randomness, Experiment repeated 10 times. In this paper, DNN's pre-training initialization parameters are shown in Table 3.

**Tab.3.** DNN model parameters

|  |  |  |  |
| --- | --- | --- | --- |
| Training parameter |  |  |  |
| Hidden layers | 6 | 4 | 3 |
| Number of neurons | 500/250/100/50/20 | 500/200/50/20 | 500/200/50 |
| Max number of epochs | 1000 | 800 | 800 |
| Learning rate | 0.01 | 0.02 | 0.01 |

DNN training uses a stochastic gradient descent method, and the maximum number of iterations of DNN in each layer is 1000,800,800 times respectively. Compared with the simulation results without feature fusion, the proposed method is validated. It can be seen from Table 5 that the diagnostic accuracy of the proposed method in frequency class and other faults is higher than that of ordinary neural networks and traditional deep neural networks DNN.

**Tab.6.** The classification results by the proposed model

|  |  |
| --- | --- |
| Method | Diagnostic accuracy |
| GFFDNN | 98.96 |
| DNN | 91.52 |
| BP | 85.56 |

It can be seen from Table 6 that the diagnostic accuracy of the proposed method is higher than that of the traditional deep neural network method and the BP algorithm, which proves that the slope curvature can increase the frequency characteristics of the data to a certain extent. Make the diagnosis better.

**5 结论**

在本文中，针对频率类故障在时域中难以有效诊断的问题，提出了一种基于DNN的几何特征融合的故障在线诊断方法。其主要创新点是通过引入原始数据的几何特征解决了当频率类故障存在的幅值相近且频繁过零点带来时域检测不准确的问题。将原始、斜率、曲率数据分别构建自动编码器模型，这样可以充分挖掘数据隐含的频率信息。然后将得到的数据特征进行特征融合，其特征维度增大使得分类模型训练的更准确。此外，该方法解决了时域故障分类困难的问题，使得频率类故障在时域中可以有效的检测出来，这就使得工业中更具有实用价值的实时故障诊断得以实现，这将有助于机械设备的健康状态检测。滚动轴承实验平台是验证了所出提方法的有效性。

**参考文献：**

[1] D. H. Zhou, Y. Liu and X. He, "Review on fault classification techniques for closed-loop systems," *Acta Automatica Sinica*, vol.39, no. 11, pp. 1933-1943,2013.

[2] F. N. Zhou, J. H. Park and Y. J. Liu, "Differential feature based hierarchical PCA fault detection method for dynamic fault," *Neurocomputing* ,vol.202, pp. 27-35, 2016.

[3] C. Li, RV. Sanchez, G. Zurita, M. Cerrada, D. Cabrera and R. E. Vásquez , "Multimodal deep support vector classification with homologous features and its application to gearbox fault classification," *Neurocomputing*, vol.168, pp.119-127, 2015.

[4] X. He, Z. Wang, Y. Liu and D. H. Zhou, "Least-Squares Fault Detection and Classification for Networked Sensing Systems Using A Direct State Estimation Approach," in *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1670-1679, Aug. 2013.

[5] D. Zhao, D. Shen and Y. Q. Wang, "Fault classification and compensation for two-dimensional discrete time systems with sensor faults and time-varying delays, " *Int. J. Robust Nonlinear Control*, doi: [10.1002/rnc.3742](http://dx.doi.org/10.1002/rnc.3742" \t "_blank" \o "Link to external resource: 10.1002/rnc.3742), 2017.

[6] H. Li and D.Y. Xiao, "Surver on data driven fault classification methods", *Control and Decision*, vol. 26, no. 1, pp. 1-9+16, 2011.

[7] R. M. An and Y. Gao, "Spacecraft fault classification based on hierarchical neural network," *Spacecraft environment engineering*, vol. 30, no. 2, pp. 203-208, 2013.

[8] F. N. Zhou , C. L. Wen, Y. B. Leng and Z. G. Chen, "A data-driven fault propagation analysis method", *Journal of Chemical Industry and Engineering(China)*, vol. 61, no. 8, pp. 1993-2001,2010.

[9] H. Q. Ji, X. He and D. H. Zhou, "On the use of reconstruction-based contribution for fault classification, "*Journal of Process Control*, vol. 40, pp. 24-34, 2016.

[10] D. J. Yu, M. F. Chen, J. S. Cheng and Y. Yang, "A fault classification approach for rotor systems based on empirical mode decomposition method and support vector machines, " *Proceedings of the Chinese society for electrical engineering*, vol. 26,no. 16, pp. 162-167, 2006.

[11] M. Gan, C. Wang and C. A. Zhu, "Construction of hierarchical classification network based on deep learning and its application in the fault pattern recognition of rolling element bearings" *Mechanical Systems and Signal Processing*, vol.72-73, pp. 92–104, 2016.

[12] G. F. Bin, J. J. Gao, [X. J. Li](http://www.sciencedirect.com/science/article/pii/S0888327011003207) and [B. S. Dhillon](http://www.sciencedirect.com/science/article/pii/S0888327011003207), "Early fault classification of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network," *Mechanical Systems and Signal Processing*, vol. 27, pp. 696–711,2012.

[12] A. Widodo and BS. Yang, "Support vector machine in machine condition monitoring and fault classification," *Mechanical Systems and Signal Processing*, vol. 21,no. 6, pp.2560-2574, 2007.

[13] D. Zhao, Z. P. Lin and Y. Q. Wang, "Integrated state/disturbance observers for two-dimensional linear systems," in *IET Control Theory & Applications*, vol. 9, no. 9, pp. 1373-1383, 2015.

[15] Q. Hu, Z. J. He, Z. S. Zhang and Y. Y. Zi , "Fault classification of rotating machinery based on improved wavelet package transform and SVMs ensemble ," *Mechanical Systems and Signal Processing* ,vol. 21, no.2, pp. 688–705, 2007.

[16] L. Y. Wang, W. G. Zhao, Y. Liu, "Rolling Bearing Fault Classification Based on Wavelet Packet- Neural Network Characteristic Entropy," Advanced Materials Research, Vols. 108-111, pp. 1075-1079, 2010.

[17] Y. Yang and W. Tang, "Study of remote bearing fault classification based on BP Neural Network combination," *2011 Seventh International Conference on Natural Computation*, pp. 618-621,Shanghai, 2011.

[18] L. Jiang, Q. Li, J. Cui and J. Xi, "Rolling bearing fault classification based on higher-order cumulants and BP neural network," *The 27th Chinese Control and Decision Conference (2015 CCDC)*, pp. 2664-2667, Qingdao, 2015.

[19] T. Kuremoto, S. Kimura, K. Kobayashi and M. Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," *Neurocomputing*, vol. 137, pp. 47–56, 2014.

[20] S. M. Zhang, F. L. Wang, S. Tan and S. Wang, "A fully automatic onine mode identiflcation method for multi-mode processes," *Acta Automatica Sinica*, vol. 42, no.1, pp.60-80, 2016.

[21] Jürgen Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.

[22] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol.313, pp. 504-507, 2006.

[23] H. Ze, A. Senior and M. Schuster, "Statistical parametric speech synthesis using deep neural networks," *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 7962-7966, Vancouver, BC, 2013.

[24] B. Song, S. Tan and H. B. Shi, "Key principal components with recursive local outlier factor for multimode chemical process monitoring," *Journal of Process Control*, vol.47,pp.136–149, 2016.

[25] L. P. Zhao, C. H. Zhao and F. R. Gao, "Inter-batch-evolution-traced process monitoring based on inter-batch mode partition for multiphase batch processes," *Chemometrics and Intelligent Laboratory Systems*, vol.138, pp.178–192, 2014.

[26] Y. W. Zhang, C. Wang and R. Q. Lu, "Modeling and monitoring of multimode process based on subspace separation," *Chemical Engineering Research and Design*, vol.91, no.5, pp.831–842, 2013.

[27] F. N. Zhou, C. L. Wen, T. H. Tang and Z. G. Chen, "DCA based multiple faults classification method," *Acta Automatica Sinica*, vol.35, no.7, pp. 971-982, 2009.

[28] P. Tamilselvan and P. F.Wang, "Failure classification using deep belief learning based health state classification," *Reliability Engineering and System Safety*, vol. 115, pp. 124–135, 2013.

[29] R. B. Huang, C. Liu, G. Q. Li, and J. L. Zhou, "Adaptive Deep Supervised Autoencoder Based Image Reconstruction for Face Recognition," Mathematical Problems in Engineering, vol. 2016, Article ID 6795352, 14 pages, 2016.

[30] H. M. Liu, L. F. Li, and J. Ma, "Rolling Bearing Fault Classification Based on STFT-Deep Learning and Sound Signals," Shock and Vibration, vol. 2016, Article ID 6127479, 12 pages, 2016.

[31] P. L. Wang , C. J. Xia, "Fault detection and self-learning identification based on PCA-PDBNs," *Chinese Journal of Scientific Instrument*, vol.36, no. 5, pp.1147-1154, 2015.

[32] R. Pang, Z. B. Yu,  W. Y. Xiong and H. Li , "Faults recognition of high -speed train bogie based on deep learning ," *Journal of Railway Science and Engineering*, vol.12, no.6, pp. 1283-1288, 2015.

[33] C. Lu，Z.Y. Wang, W. L. Qin and J. Ma, "Fault classification of rotary machinery components using a stacked denoising autoencoder-based health state identification," *Signal Processing*, vol.130, pp.377-388, 2017.

[34] F. Jia, Y. G. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent classification of rotating machinery with massive data," *Mechanical Systems and Signal Processing*, vol.72-73, pp. 303–315, 2016.

[35] Bearing data Centre, Case Western Reserve University, Available:

*<http://csegroups.case.edu/bearingdatacenter/home>*