

# 基于 EDBN-SVM 的高速列车故障分析

郭超<sup>1</sup> 杨燕<sup>1</sup> 金炜东<sup>2</sup>

(西南交通大学信息科学与技术学院 成都 611756)<sup>1</sup> (西南交通大学电气工程学院 成都 611756)<sup>2</sup>

**摘要** 深度学习作为机器学习领域的新热点,为故障诊断技术领域研究开拓了新的思路。针对高速列车进行故障分析的重要性,将深度学习和集成学习相结合,提出一种基于 EDBN-SVM(Ensemble Deep Belief Network-Support Vector Machine)的故障诊断模型。首先对高速列车振动信号进行快速傅立叶变换,其次分析确定了 EDBN-SVM 模型的参数,然后将信号的 FFT 系数作为 EDBN-SVM 模型的可视层输入,并逐层学习高层特征,最后利用多个 SVM 分类器进行识别并对识别结果进行集成。为评估该方法的有效性,采用实验室数据和仿真数据进行实验测试,并与传统的几种故障分析方法进行对比。结果表明,该方法的故障识别效果优于传统的故障分析方法,同时稳定性更好。

**关键词** 高速列车,故障分析,快速傅立叶变换,深度信念网络

**中图分类号** TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.12.052

## Fault Analysis of High Speed Train Based on EDBN-SVM

GUO Chao<sup>1</sup> YANG Yan<sup>1</sup> JIN Wei-dong<sup>2</sup>

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China)<sup>1</sup>

(School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China)<sup>2</sup>

**Abstract** As a new hot spot in the field of machine learning, deep learning has opened up new ideas for the research of fault diagnosis. In view of significance of fault analysis for high speed train, combining deep learning and ensemble learning, a new fault diagnosis model based on EDBN-SVM(Ensemble Deep Belief Network-Support Vector Machine) was proposed. Firstly, we preprocessed the vibration signal of high speed train by fast fourier transform (FFT). Secondly, we analyzed the parameters of the EDBN-SVM model, then we set the FFT coefficients as the input of the visible layer of EDBN-SVM model, and used the model to learn high-level features layer by layer. Finally, we utilized multiple SVM classifiers to recognize faults, and combined the recognition results. In order to evaluate the validity of this method, we selected the laboratory data and the simulation data to conduct experiments, and compared it with the traditional fault analysis methods. The results show that the fault recognition effect and the stability of this method are better than traditional methods.

**Keywords** High speed train, Fault analysis, Fast Fourier transform, Deep belief network

## 1 引言

近年来,中国高速铁路飞速发展,然而高速铁路存在的安全隐患不容忽视。走行部是高速列车的重要组成部分,对列车的安全性和舒适度起到重要作用<sup>[1]</sup>。列车走行部存在空气弹簧故障、横向减振器故障、抗蛇行减振器故障 3 种典型故障。空气弹簧起到减弱列车垂向振动和冲击的作用,发生空气弹簧失气会导致列车垂向振动加剧,降低乘客舒适度。横向减振器用于减小路面不平顺产生的列车横向振动,横向减振器失效可能导致列车发生侧滚。抗蛇形减振器用来抑制走行部蛇形运动,一旦出现故障可能导致列车蛇形运动失稳现象,严重影响列车的安全运行。因此,研究高速列车走行部的故障诊断和性能分析方法,对保障高速列车安全运行具有重要意义。通过安装在高速列车上的传感器来采集反映高速列车走行部运行状况的振动信号,是当前监测高速列车的运行

状态的主要技术之一。

由于机械故障信号具有非线性、非平稳特性,通常采用时频域分析的方法分析故障信号。黄采伦等人利用离散余弦变换和小波变换方法,对机车走行部的牵引齿轮故障信号进行诊断<sup>[2]</sup>。Lei 等人提出基于聚合经验模态分解(EEMD)的方法对旋转机故障信号进行特征提取,然后进行故障诊断<sup>[3]</sup>。Qin 等采用小波熵表示信号特征,然后利用支持向量机(SVM)进行识别<sup>[4]</sup>。陈云凤等采用基尼系数、统计假说等 4 种方法进行特征选择,最后采用聚类集成算法进行故障识别<sup>[5]</sup>。Li 基于 PCA 和自组织映射(SOM)神经网络实现了滚动轴承故障诊断<sup>[6]</sup>。赵晶晶等人提出了一种基于经验模态分解(EMD)和模糊熵结合的特征提取方法,然后使用 BP 神经网络对高速列车进行故障识别<sup>[7]</sup>。然而采用 EMD 等特征提取方式并不能保证提取出稳定的特征。同时面对故障诊断问题,我们并不能确切知道哪些特征更好地反映出故障信号的

到稿日期:2015-10-11 返修日期:2016-01-22 本文受国家自然科学基金项目(61134002,61572407)资助。

郭超(1992-),男,硕士,CCF 会员,主要研究方向为智能信息处理;杨燕(1964-),女,博士,教授,主要研究方向为数据挖掘、集成学习等;金炜东(1959-),男,博士,教授,主要研究方向为智能信息处理、模式识别等。

特征。而且采用 SVM 和 BP 神经网络等浅层结构的分类器时,对于复杂的分类问题,其泛化能力会受到一定限制<sup>[8]</sup>。

深度学习是机器学习的一个新领域,它掀起了机器学习的第二次浪潮<sup>[9]</sup>。深度学习的发展避免了进行大量人工提取数据特征的工作,使分类模型更加智能化。深度信念网络(Deep Belief Network, DBN)<sup>[10]</sup>是一个包含多个隐藏层的神经网络模型。DBN 使用多个受限玻尔兹曼机(RBM)<sup>[11]</sup>,通过逐层无监督贪婪训练方法使网络学习到更高层的特征。与 BP 神经网络等浅层分类器相比,DBN 可以完成复杂函数的非线性逼近,并避免过拟合等问题出现。在 DBN 提出的短短数年里,它已经广泛地运用在语音识别<sup>[12]</sup>、目标识别<sup>[13]</sup>、故障诊断<sup>[14]</sup>以及手写体识别<sup>[15]</sup>等领域。但 DBN 的输出具有随机性,性能不稳定。集成学习通过使用多个学习器解决同一个问题,可以更有效地提高学习系统的泛化能力和可靠性。Hansen 和 Salamon<sup>[16]</sup>开创性地将多个神经网络进行集成,显著提高了整个神经网络的性能。目前,集成学习已经在客户资源管理<sup>[17]</sup>、文本分类<sup>[18]</sup>等领域成功运用。

本文通过分析传感器采集的高铁振动信号的时域和频域特性,结合深度学习和集成学习,提出了一种基于 EDBN-SVM 的模型对高速铁路走行部进行故障分析。提取振动信号的快速傅立叶变换(FFT)系数作为 EDBN-SVM 的可视层输入,通过逐层无监督训练每个 RBM,对提取出的振动信号高层特征进行分析。然后对 EDBN-SVM 模型网络层数、隐藏层单元个数及 SVM 个数 3 个参数进行实验分析。最后,使用确定参数后的 EDBN-SVM 模型进行故障分析。

## 2 相关原理

### 2.1 RBM 原理

受限玻尔兹曼机(RBM)是一个基于能量的无向生成模型。如图 1 所示,它是由可视层和隐藏层两层组成,可视单元  $v$  和隐藏单元  $h$  之间通过权值  $w$  对称双向全连接,而相同层内的单元之间无连接。由于 RBM 的这种特殊结构,对于伯努利-伯努利 RBM,隐藏单元  $h_j=1$  的条件概率为:

$$p(h_j=1|v) = \frac{1}{1 + \exp(-b_j - \sum_{i=1}^n W_{ij}v_i)} \quad (1)$$

可见单元  $v_i=1$  的条件概率为:

$$p(v_i=1|h) = \frac{1}{1 + \exp(-a_i - \sum_{j=1}^m W_{ij}h_j)} \quad (2)$$

其中,  $a_i$  表示可见单元  $i$  的偏置;  $b_j$  表示隐藏单元  $j$  的偏置;  $w_{ij}$  为可见单元  $i$  和隐藏单元  $j$  之间的连接权值;  $n$  和  $m$  分别为可视单元和隐单元的个数;  $v_i, h_j \in \{0, 1\}$  分别表示可见单元  $i$  和隐藏单元  $j$  的状态。

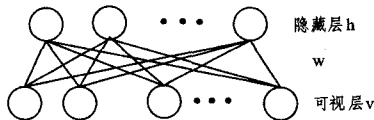


图 1 受限玻尔兹曼机模型

当给定输入  $v$  和随机初始化的网络权值  $w$  时, RBM 网络的系统能量为:

$$E(v, h) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i h_j w_{ij} \quad (3)$$

基于该能量函数,可得到  $(v, h)$  的联合概率分布:

$$P(v, h) = \frac{e^{-E(v, h)}}{\sum_v \sum_h e^{-E(v, h)}} \quad (4)$$

同样,对于高斯-伯努利 RBM,其条件概率为:

$$p(h|v) = \frac{1}{1 + \exp(-b_j - \sum_{i=1}^n W_{ij}v_i)} \quad (5)$$

$$p(v|h) = N(a_i + \sum_{j=1}^m W_{ij}h_j, 1) \quad (6)$$

高斯-伯努利 RBM 网络的系统能量为:

$$E(v, h) = -\frac{1}{2} \sum_{i=1}^n (v_i - a_i)^2 - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i h_j w_{ij} \quad (7)$$

其中,  $v$  服从均值为  $a_i + \sum_{j=1}^m W_{ij}h_j$ 、方差为 1 的高斯分布。通过训练高斯-伯努利 RBM 可将实值随机变量转换为二进制随机变量,然后使用伯努利-伯努利 RBM 对其进行训练。

学习 RBM 的主要任务是求出网络参数  $\theta = \{w, a, b\}$ , Hinton 提出利用对比散度(Contrastive Divergence, CD)算法<sup>[19]</sup>求出联合概率分布  $p(v, h)$  的最大值从而求得  $\theta$  的最优值。CD 算法通过计算两个概率分布之间的最小 KL 距离以求取  $p(v, h)$  最大值:

$$CD_n = KL(p_0 \| p_\infty) - KL(p_n \| p_\infty) \quad (8)$$

其中,  $p_0$  为 RBM 初始状态的概率分布;  $p_n$  为经过  $n$  步吉布斯采样后 RBM 的概率分布;  $p_\infty$  为 RBM 趋于稳定后的概率分布。文献<sup>[20]</sup>经实验证明, CD 算法不断地将  $p_n$  赋值给  $p_0$ , 经过多次求斜率修正参数  $\theta$ ,  $CD_n$  趋向于 0, 即  $p_n$  趋于  $p_\infty$ 。通常取  $n=1$  即可满足计算需要, 参数  $\theta$  的更新规则为:

$$\begin{aligned} W_t &= W_{t-1} + \eta(E_{data}(v_i h_j) - E_{recon}(v_i h_j)) \\ b_t &= b_{t-1} + \eta(E_{data}(v_i) - E_{recon}(v_i)) \\ c_t &= c_{t-1} + \eta(E_{data}(h_j) - E_{recon}(h_j)) \end{aligned} \quad (9)$$

其中,  $\eta$  为学习率;  $E_{data}(\cdot)$  为观测数据概率分布决定的期望;  $E_{recon}(\cdot)$  是 RBM 模型重构后的概率分布决定的期望。图 2 给出了基于 CD 算法的 RBM 训练过程。  $v_0$  是可视层输入,  $h_0$  是由式(1)得到的隐藏层特征向量, 利用式(2)得到重建后的特征  $v_1$ , 然后由  $v_1$  推导出隐藏层特征向量  $h_1$ 。最后由式(9)修正参数  $\theta = \{w, a, b\}$ 。

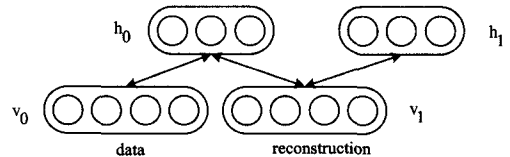


图 2 基于 CD 算法的 RBM 训练过程

### 2.2 DBN 原理

深度信念网络(DBN)是广泛研究和应用的深度学习模型。DBN 模型由若干层 RBM 堆叠组成。图 3 展示了包含两层 RBM 的 DBN 高层特征提取模型。DBN 模型的训练过程主要分为无监督预训练和有监督微调两步。DBN 模型首先利用贪婪算法无监督地单独训练每一个 RBM, 当底层 RBM 训练完成, 将其隐藏层的输出作为高层 RBM 的输入, 从而学习更高层特征。然而, 每层 RBM 只能保障自身的权值对该层特征向量映射最优, 无法确保对整个 DBN 特征向量映射取得最优。所以, 当每层 RBM 训练完成后, 利用 BP 算法将误差从上到下反向传播到每一层 RBM, 并微调每层 RBM 之间的参数, 从而使 DBN 模型参数达到最优。相对 BP 神经网络, DBN 模型有效地克服了因随机初始化权值而容易陷入局部最优和训练时间较长的缺点。

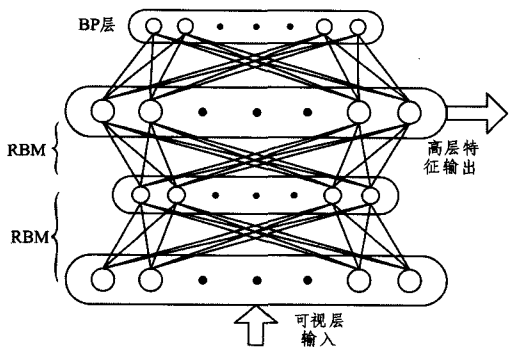


图3 DBN高层特征提取模型

DBN网络训练完成后,利用网络参数将可视层输入特征映射到高层,利用式(10)提取高层特征。

$$y = f(v; \theta_{DBN}) \quad (10)$$

其中,  $v$  表示可视层输入特征,  $\theta_{DBN}$  为 DBN 的网络参数,  $y$  为经过 DBN 网络映射后提取的高层特征。

### 2.3 集成原理

多数投票法(Majority Voting)是一种经典的数据集成方法,对组内决策简单且有效<sup>[21]</sup>。它利用多个分类器单独对样本进行分类,并将该样本分到大多数分类器所属类别。多数投票法的输出策略为:

$$\hat{y} = \underset{j=1,2,\dots,k}{\operatorname{argmax}} \{i | f_i(x) = d_j\} \quad (11)$$

其中,  $f_i(\cdot)$  ( $i=1, 2, \dots, k$ ) 为第  $i$  个分类器的判决函数,  $d_j$  表示将该样本划分到第  $j$  类。给定测试样本  $x$ , 每个分类器将其分到第  $j$  类的个数为  $m_j = \{i | f_i(x) = d_j\}$ 。

集成学习通过将多个分类器组合在一起构成高精度分类器,从而避免单个分类器的不稳定性和低精度。支持向量机(SVM)是一种广泛应用的二分类器, SVM 分类器实现多分类时,通常采用一对一、一对多等组合策略。但单个 SVM 分类器的性能不稳定,会出现泛化能力差等问题。比如采用一对多组合方式时,可能出现误差无界。为解决此类问题,本文采用多数投票集成策略,将多个 SVM 多分类器进行集成,并与 DBN 网络结合,构造 EDBN-SVM 高速列车故障诊断模型。

### 2.4 EDBN-SVM 故障诊断模型

基于 EDBN-SVM 的高速列车故障诊断模型如图 4 所示,其主要包含 3 个步骤:

- (1)在预处理阶段对高速列车振动信号分段并进行 FFT,将得到的 FFT 系数进行归一化处理,并把归一化后的数据作为 DBN 的底层输入。
- (2)利用 2.2 节的 DBN 高层特征提取模型逐层进行无监督训练,并利用 BP 算法进行微调,最后提取出高层特征。
- (3)把学习到的高层特征输入到多个 SVM 分类器进行集成学习,得到故障识别率。

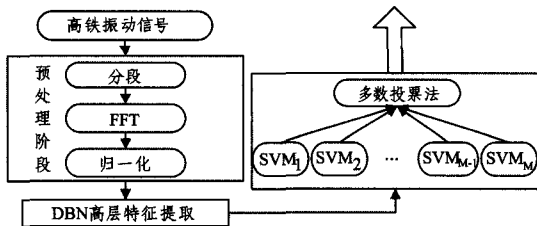


图4 EDBN-SVM 高速列车故障诊断模型

EDBN-SVM 故障诊断模型具体算法如下。

输入:高速列车振动信号样本  $f$

输出:故障识别准确率 Acc 网络层数  $N$ , 学习率  $\eta$ , 第  $i$  层 RBM 权值矩阵  $w_i$ , 第  $i$  个隐藏层单元向量  $h_i$ , 第  $i$  个输入偏置  $a_i$ , 第  $i$  个输出偏置  $b_i$ , DBN 网络提取的高层特征  $y$

#### 1. 高层特征提取过程:

- (1)对信号  $f$  进行 FFT, 将其 FFT 系数归一化到  $[0, 1]$ , 记为  $v_0$ , 将  $v_0$  作为 DBN 可视层输入;
- (2)随机初始化  $\theta_i = \{w_i, a_i, b_i\}$ ;
- (3)对输入特征向量  $v_0$ , 由式(1)和式(2)计算特征向量  $h_{0i}, v_{1i}, h_{1i}$ ;
- (4)利用式(9)更新参数  $w_i, a_i, b_i$ ;
- (5)每层 RBM 训练完成后, 令  $v_i = h_{i+1}$ ;
- (6)重复步骤(2)~步骤(5)  $N$  次, 完成所有 RBM 的训练;
- (7)使用 BP 算法反向传播错误信息, 微调整个 DBN 网络;
- (8)整个 DBN 网络训练完成后, 利用式(10)将可视层特征  $f$  映射到高层特征  $y$ 。

#### 2. 识别过程:

- (1)将 DBN 网络学习到的高层特征  $y$  输入到多个 SVM 多分类器;
- (2)采用多数投票法, 利用式(11)对每个分类器的识别结果进行集成, 并计算故障识别准确率 Acc。

## 3 故障分析仿真实验

### 3.1 高速列车振动信号分析

为评估本文提出的 EDBN-SVM 故障诊断模型的性能, 采用某实验室采集的某型号列车实验室数据和故障仿真数据进行实验。该实验室数据记录了列车 4 种工况下的振动信号, 在不同工况下测试 40~440km/h 之间不同速度下的振动信号, 直至列车出现失稳。一路振动信号包含 14616 个数据点, 采样频率为 243Hz。本文将每种速度下的实验室数据划分为 56 段, 每段包含 260 个样本点。仿真数据由仿真系统(SIMPACK 和轨道谱)生成, 对某型号动车组中动车走行部故障进行实验仿真。一路仿真数据有 51000 个数据点, 采样频率为 243Hz。本文将每种速度下的仿真数据划分为 180 段, 每段包含 260 个点。

图 5 为速度为 140km/h 时, 4 种工况下列车振动信号时域图及对应的频谱图。

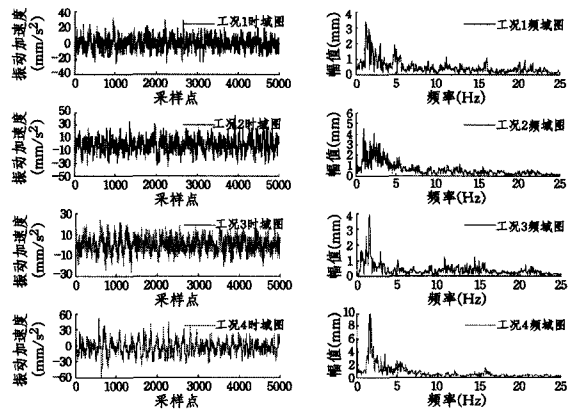


图5 4种工况下列车振动信号时频图

其中, 工况 1 为正常状态; 工况 2 为空气弹簧失气; 工况 3 为横向减振器失效; 工况 4 为抗蛇形减振器故障。从工况 2 时频图中可以看出, 当发生空气弹簧失气时, 列车振动会加剧, 垂向振动加速度增大, 对应在 3Hz 附近的频域振幅明显增大。工况 3 时频图显示横向减振器失效导致在其固有特征频率约 2.5Hz 处的振动频率增大。出现抗蛇形减振器故障时, 会引起列车蛇形运动失稳, 时域波形的振幅也会增大, 同时频域振幅增加, 故障特征集中在 1.5Hz 左右。同时可以看

出几种工况的时域特征较为相似,很难进行区分。而频域特征差异较为明显,并且特征都集中在6Hz以内。因此实验取6Hz内频谱(100个采样点)作为EDBN-SVM模型的可视层输入。

### 3.2 基于EDBN-SVM的高层特征提取

实验选取速度为160km/h时的实验室数据和仿真数据进行测试,从而验证高速列车振动信号使用EDBN-SVM模型的特征提取性能。DBN网络使用3个RBM,网络结构设置为100-150-100-2,每个RBM学习率设为0.1,迭代训练70次。图6(a)为实验室数据经过训练得到的高层特征分布图。从中可以看出工况2和工况3具有较好的类间分离性和类内聚集性,工况1和工况4有少量混叠现象发生。图6(b)为仿真数据经过EDBN-SVM模型提取的高层特征分布图。图中工况2和工况4聚集性较好同时可以较好地分离,工况1和工况3有部分特征发生混叠,但总体上高层特征分离较好。

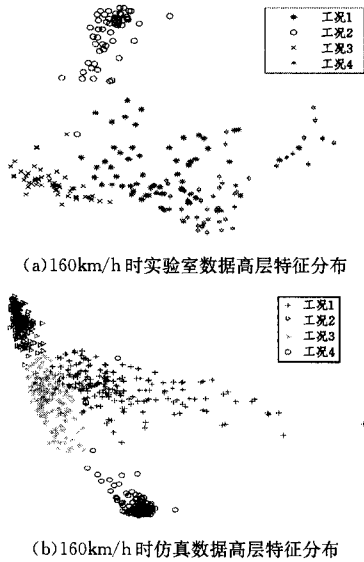


图6 经EDBN-SVM模型训练的高层特征分布图

### 3.3 基于EDBN-SVM的故障分析性能

为分析EDBN-SVM网络参数设置对实验结果的影响,本实验将研究不同网络层数、不同隐藏层单元个数及不同的SVM个数对故障诊断性能的影响。RBM学习速率设为0.1,迭代训练60次。首先研究不同网络层数的影响,将网络每个隐藏层单元数、输出层单元数均设为100,使用1个SVM分类器进行故障识别。图7显示了3-10层网络对实验室数据的故障识别率。从图中可以看出,EDBN-SVM模型网络层数为3时,不同速度下的故障识别率均相对较高。

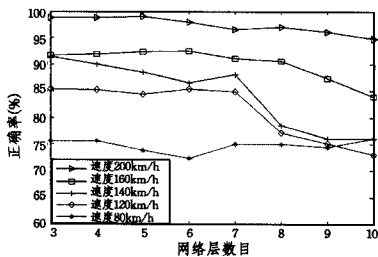


图7 不同网络层数下的故障识别率

进一步使用具有3层网络、1个SVM分类器的模型,设置不同的隐藏层单元数进行实验。5种速度下不同隐藏层单元数的故障识别率如图8所示。结果表明,隐藏层单元数在

250以内时,各速度下的故障识别率变化很小;仅在隐藏层单元数为300,速度为120km/h时的识别率较差;在隐藏层单元数为150,不同速度时的故障识别率相对较高。

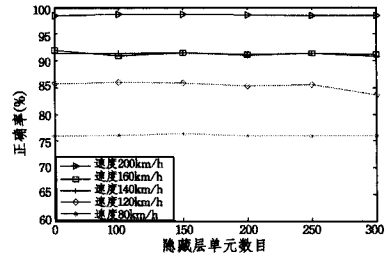


图8 不同隐藏层单元个数下的故障识别率

为研究不同的SVM个数对实验结果的影响,实验使用100-150-100的网络进行训练,将SVM个数 $M$ 分别设置为1-15之间进行仿真(为尽量减小其时间复杂性,SVM个数设为15以内)。图9(a)展示了不同SVM个数的故障识别率;图9(b)展示了其对应的均方根误差(RMSE)曲线。RMSE反映了测量的精密度,文献[22]给出了RMSE的定义。从图9(a)可以看出,当 $M$ 取9时,不同速度下的故障识别率都较高,并且识别率的平均值达到了最高。同时,图9(b)示出的RMSE曲线表明, $M=9$ 时对应的RMSE较小,并且其均值也达到最小。综合考虑其故障识别率和均方根误差,将EDBN-SVM模型参数 $M$ 设置为9。

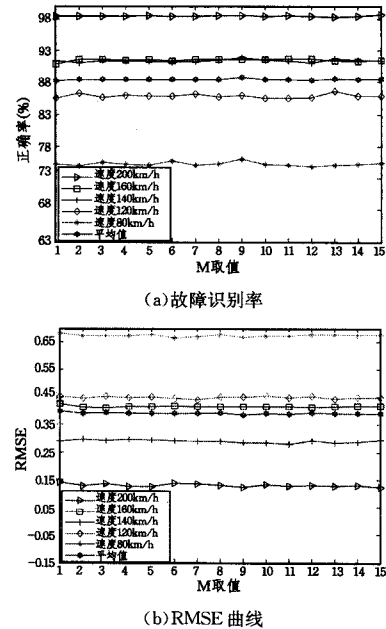


图9 不同SVM个数的实验结果图

### 3.4 实验结果与分析

为验证本文提出的故障诊断模型的性能,实验采用传统的故障分析方法进行对比。文献[3,7,23]提出利用经验模态分解(EMD)方式对机械振动信号分解并提取近似熵或模糊熵,然后使用BP神经网络、SVM分类器进行故障识别。文献[14]对信号的FFT系数直接使用BP和SVM两种浅层结构分类器进行识别,对应方法为FFT-BP和FFT-SVM。实验将DBN-SVM和EDBN-SVM的网络结构设置为100-150-100,SVM个数分别设置为1和9,使用实验室数据分别对几种故障分析方法的性能进行测试。

表1列出了10种方法的故障诊断识别率和标准差,其中

AE代表近似熵,FE代表模糊熵。对比10种故障分析方法可以看出,采用深度学习方法 EDBN-SVM 和 DBN-SVM 的故障诊断效果远好于传统的故障分析方法。基于 EDBN-SVM 的故障分析方法对工况 1 和工况 4 的识别率最高,对工况 3 的识别率达到次高值,并且其平均识别率最高。同时,使用 EDBN-SVM 模型进行故障诊断的标准差也相对较小。

表 1 实验室数据故障诊断结果

故障分析方法	工况 1 识别率 (%)	工况 2 识别率 (%)	工况 3 识别率 (%)	工况 4 识别率 (%)	平均识别率 (%)	标准差
EMD-AE-BP	24.59	99.88	78.59	59.41	65.62	0.0394
EMD-FE-BP	18.47	99.88	75.76	62.35	64.12	0.0501
EMD-FE-AE-BP	15.65	100	81.76	60.59	64.50	0.0441
EMD-AE-SVM	8.47	93.41	85.41	14.35	50.41	0.0316
EMD-FE-SVM	2	99.41	98.24	15.53	53.79	0.0223
EMD-FE-AE-SVM	24.24	100	90.47	43.53	64.56	0.0363
FFT-BP	50.57	82.35	76.21	57.13	66.56	0.1160
FFT-SVM	51.53	97.20	93.48	60.89	75.78	0.0377
DBN-SVM	76.87	98.26	96.12	81.91	88.29	0.0312
EDBN-SVM	77.86	98.57	96.71	83.25	89.10	0.0288

采用仿真数据集进行相同的测试,并与 FFT-SVM,FFT-BP 以及 DBN-SVM 故障分析方法进行对比。从表 2 列出的故障诊断结果可以看出,EDBN-SVM 模型对 4 种工况的故障识别率最高,平均识别率达到 90.01%,并且标准差取得最小值 0.0131。这验证了本文提出的 EDBN-SVM 模型具有良好的故障诊断性能,同时具有较高的稳定性。

表 2 仿真数据故障诊断结果

故障分析方法	工况 1 识别率 (%)	工况 2 识别率 (%)	工况 3 识别率 (%)	工况 4 识别率 (%)	平均识别率 (%)	标准差
FFT-BP	61.37	89.89	68.63	91.03	77.73	0.0770
FFT-SVM	75.53	94.81	80.01	98.47	87.20	0.0156
DBN-SVM	81.39	96.70	82.36	98.42	89.72	0.0168
EDBN-SVM	81.82	96.82	82.85	98.56	90.01	0.0131

**结束语** 本文将深度学习与集成学习相结合,提出了一种 EDBN-SVM 模型用于高速列车的故障分析。首先研究了不同工况列车振动信号的时频特性,分析得出列车的频域特征可以更好地反映出列车振动信号的特征,并且故障特征主要集中在 6Hz 以内;然后展示了 EDBN-SVM 模型良好的特征提取能力;接下来分析了该深度学习模型的参数(网络层数、隐藏层单元个数及 SVM 个数)对故障识别能力的影响。结果显示,具有 100-150-100 网络结构、SVM 多分类器个数为 9 的 EDBN-SVM 模型表现出的性能最优。最后,分别使用实验室数据和仿真数据测试所提模型的故障识别率,并与 EMD-FE-BP 等传统的故障分析方法进行对比。实验结果表明,EDBN-SVM 故障诊断模型可以有效识别出高速列车的 4 种典型工况,并且在识别率和稳定性方面均优于传统的故障分析方法。集成学习的引入增强了该深度学习模型的容错性,并且在一定程度上提高了其准确率和稳定性。然而,在实际高速列车运行过程中,几种典型的故障可能会同时发生。本文仅对单一故障进行了分析,今后将进一步研究多种故障混合工况下的故障分析方法。

### 参考文献

[1] Garcia C, Lehner A, Strang T, et al. Comparison of collision avoidance systems and applicability to rail transport[C]//Proceedings of the 7th International Conference on Intelligent

Transport Systems. France; Sophia Antipolis, 2007: 1-6

[2] Huang Cai-lun, Fan Xiao-ping, Chen Chun-yang, et al. Fault diagnosis method of locomotive driven gear based on envelopment analysis of wavelet coefficients extraction and DCT[J]. Journal of the China Railway Society, 2008, 30(2): 98-102 (in Chinese)

黄采伦,樊晓平,陈春阳,等.基于小波系数提取及离散余弦包络分析的机车牵引齿轮故障诊断方法[J].铁道学报,2008,30(2): 98-102

[3] Lei Ya-guo, He Zheng-jia, Zi Yan-yang. Application of the EEMD method to rotor fault diagnosis of rotating machinery[J]. Mechanical Systems and Signal Processing, 2009, 23(4): 1327-1338

[4] Qin Na, Jin Wei-dong, Huang Jin, et al. High speed train bogie fault signal analysis based on wavelet entropy feature[J]. Advanced Materials Research, 2013, 753(12): 2286-2289

[5] Chen Yun-feng, Wang Hong-jun, Yang Yan. Fault diagnosis of high-speed rail based on clustering ensemble[J]. Computer Science, 2015, 42(6): 233-238 (in Chinese)

陈云凤,王红军,杨燕.基于聚类集成的高铁故障诊断分析[J].计算机科学,2015,42(6): 233-238

[6] Li Zhi-chun. A Simple SOM neural network based fault detection model for fault diagnosis of rolling rearings[J]. Applied Mechanics and Materials, 2013, 397: 1321-1325

[7] Zhao Jing-jing, Yang Yan, Li Tian-rui, et al. Fault diagnosis of high-speed rail based on approximate entropy and empirical mode decomposition[J]. Computer Science, 2014, 41(1): 91-94 (in Chinese)

赵晶晶,杨燕,李天瑞,等.基于近似熵及 EMD 的高铁故障诊断[J].计算机科学,2014,41(1): 91-94

[8] Sun Zhi-jun, Xue Lei, Xu Yang-ming, et al. Overview of deep learning[J]. Application Research of Computer, 2012, 29(8): 2806-2810 (in Chinese)

孙志军,薛磊,许阳明,等.深度学习研究综述[J].计算机应用研究,2012,29(8): 2806-2810

[9] Guo Li-li, Ding Shi-fei. Research progress on deep learning[J]. Computer Science, 2015, 42(5): 28-33 (in Chinese)

郭丽丽,丁世飞.深度学习研究进展[J].计算机科学,2015,42(5): 28-33

[10] Hinton G, Osindero S, Teh Y. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554

[11] Roux N L, Bengio Y. Representational power of restricted boltzmann machines and deep belief networks[J]. Neural Computation, 2008, 20(6): 1631-1649

[12] Mohamed A, Dahl G, Hinton G. Acoustic modeling using deep belief networks[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2012, 20(1): 14-22

[13] Ranzato M, Huang F, Boureau Y, et al. Unsupervised learning of invariant feature hierarchies with applications to object recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA; IEEE, 2007: 1-8

[14] Xie Ji-peng, Yang Yan, Li Tian-rui, et al. Learning features from high speed train vibration signals with deep belief networks [C]//Proceedings of the International Joint Conference on Neural Networks (IJCNN). USA; IEEE, 2014: 2205-2210

[15] Lee H, Grosse R, Ranganath R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations[C]//Proceedings of the International Conference on Machine Learning. USA; ACM, 2009: 609-616

- [16] Hansen L, Salamon P. Neural network ensembles [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(10):993-1001
- [17] Pang S, Kim D, Bang S. Fraud detection using support vector machine ensemble [C] // Proceedings of the 8th International Conference on Neural Information Processing, Shanghai: Fudan University, 2001: 1344-1349
- [18] Dong Y, Han K. A comparison of several ensemble methods for text categorization [C] // Proceedings of the International Conference on Services Computing (SCC). USA: IEEE, 2004: 419-422
- [19] Hinton G. Training products of experts by minimizing contrastive divergence [J]. Neural Computer, 2002, 14(8): 1771-1800
- [20] Carreira-Perpinan M, Hinton G. On contrastive divergence learning [OL]. <http://www.doc88.com/IP-747820956400.html>
- [21] Yang Su-hong. A voting statistical method of group decision [C] // Proceedings of the International Joint Conference on Computational Sciences and Optimization (CSO 2009). USA: IEEE, 2009: 873-875
- [22] Tanaka M, Okutomi M. A Novel Inference of a Restricted Boltzmann Machine [C] // Proceedings of the 22th International Conference on Pattern Recognition (ICPR). USA: IEEE, 2014: 1526-1531
- [23] Wang Lei, Ji Guo-yi. Fault diagnosis of rotor system based on EMD fuzzy entropy and SVM [J]. Noise and Vibration Control, 2012, 32(3): 171-176 (in Chinese)  
王磊, 纪国宜. 基于 EMD 模糊熵和 SVM 的转子系统故障诊断 [J]. 噪声与振动控制, 2012, 32(3): 171-176

(上接第 272 页)

点。首先,文献[7]的着力点是构建关于事件的本体<sup>[14]</sup>,从动作、对象、时间、环境、断言、语言表现等多个方面对事件进行详细的描述;而本文的着力点是应用描述逻辑对关于传感器网络的领域知识进行描述,然后再在这些知识的基础上结合基于阈值的方法对事件进行刻画。其次,文献[7]需要使用描述逻辑 SHOQ(D),该逻辑的推理复杂度为非确定 2 次指数时间;而本文采用了描述逻辑  $\epsilon\mathcal{L}^{++}$ ,其推理复杂度为多项式时间。总的来说,在对事件的刻画方面,本文方法考虑的因素没有文献[7]的方法全面;但在计算性能方面,本文方法在理论上要优于文献[7]。在下一步的工作中,将从事件识别的效果(包括精确率和召回率)以及实际计算性能两个方面对两种方法进行定量的分析和比较。

**结束语** 本文给出了基于语义模型的传感器网络事件检测方法,使用轻量级描述逻辑  $\epsilon\mathcal{L}^{++}$  对领域知识进行描述,在一定程度上减小了逻辑推理的复杂度。通过用描述逻辑对领域知识进行建模,弱化了元数据的异构性特征,使得数据融合顺利进行。当然描述逻辑不仅仅限于传感器网络领域,它还能应用到更加广阔的空间,尤其适用于语义 Web 环境下对复杂系统进行刻画和有效推理。

本文仅通过实例来验证所提方法的有效性,下一步将通过具体的实验评估方案来检验方法的精确度、召回率、F1 值等统计学指标。此外,本文涉及的传感器数量较少,且未考虑传感器时空特性。下一步将设计具体的实验方案,验证该方法在智能家居领域的事件检测中的实际效果。

## 参 考 文 献

- [1] Xue W W, Luo Q, Wu H J. Pattern-based event detection in sensor networks [J]. Distributed and Parallel Databases, 2012, 30(1): 27-62
- [2] El Faouzi N E, Leung H, Kurian A. Data fusion in intelligent transportation systems: Progress and challenges A survey [J]. Information Fusion, 2011, 12(1): 4-10
- [3] Abadi D J, Madden S, Lindner W. Reed: Robust, efficient filtering and event detection in sensor networks [C] // Proceedings of the 31st International Conference on Very Large Data Bases. VLDB Endowment, 2005: 769-780
- [4] Xue W, Luo Q, Chen L, et al. Contour map matching for event detection in sensor networks [C] // Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data. ACM, 2006: 145-156
- [5] Wang T Y, Yu C T. Collaborative event region detection in wireless sensor networks using markov random fields [C] // 2nd International Symposium on Wireless Communication Systems, 2005. IEEE, 2005: 493-497
- [6] Wang X R, Lizier J T, Obst O, et al. Spatiotemporal anomaly detection in gas monitoring sensor networks [M] // Wireless Sensor Networks, Springer Berlin Heidelberg, 2008: 90-105
- [7] Tang Ying-ying, Liu Wei, Su Xiao-ying, et al. An Event Detection Method Based on Extended Description Logic [J]. Application Research of Computers, 2014, 31(11): 3285-3290 (in Chinese)  
唐英英, 刘炜, 苏小英, 等. 一种基于扩展描述逻辑的事件实例检测方法 [J]. 计算机应用研究, 2014, 31(11): 3285-3290
- [8] Liu Zong-tian, Huang Mei-li, Zhou Wen, et al. Event-Oriented Ontology Research [J]. Computer Science, 2009, 36(11): 189-192 (in Chinese)  
刘宗田, 黄美丽, 周文, 等. 面向事件的本体研究 [J]. 计算机科学, 2009, 36(11): 189-192
- [9] Baader F, Bauer A, Baumgartner P, et al. A novel architecture for situation awareness systems [M] // Automated Reasoning with Analytic Tableaux and Related Methods, Springer Berlin Heidelberg, 2009: 77-92
- [10] Wang Xiao-feng, Zhang Da-peng, Wang Fei, et al. Video Event Detection Based on Semantic Trajectory [J]. Chinese Journal of Computers, 2010, 33(10): 1845-1858 (in Chinese)  
王晓峰, 张大鹏, 王绯, 等. 基于语义轨迹的视频事件探测 [J]. 计算机学报, 2010, 33(10): 1845-1858
- [11] Baader F, Brandt S, Lutz C. Pushing the EL envelope further [J]. Proc of Ijcai, 2010, 26(2): 364-369
- [12] Compton M, Barnaghib P, Bermudez L, et al. The SSN Ontology of the W3C Semantic Sensor Network Incubator Group [J]. Web Semantics Science Services & Agents on the World Wide Web, 2012, 17(4): 25-32
- [13] Ruta M, Scioscia F, Loseto G, et al. A Logic-based CoAP Extension for Resource Discovery in Semantic Sensor Networks [C] // SSN. 2012: 17-32
- [14] Liu W, Xu W, Fu J, et al. An Extended Description Logic for Event Ontology [C] // Proceedings of the 5th International Conference on Advances in Grid and Pervasive Computing. Springer-Verlag, 2010: 471-481