

12-16-2020

Fault Diagnosis for Bearings of Unbalanced Data Based on Feature Generation

Minglu Fan

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China;

Wang Yan

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China;

Zhicheng Ji

Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Fault Diagnosis for Bearings of Unbalanced Data Based on Feature Generation

Abstract

Abstract: Focus on the sample imbalance and insufficiency caused by the difficulty to obtain a sufficient number of fault samples in actual production. A model for rolling bearings by combining Convolutional Neural Networks and Synthetic Oversampling is presented. *The frequency domain signals is used as the input of the model, and the features are extracted by the Convolutional Neural Network. The new features are generated by Synthetic Oversampling and the data equalization is realized. The model completes the classification by putting all of the features into the Support Vector Machine, and the fault diagnosis of the rolling bearings is carried out.* The comparison experiments results show that the method can effectively solve the problem of data imbalance.

Keywords

rolling bearing, fault diagnosis, feature generation, convolutional neural network

Recommended Citation

Fan Minglu, Wang Yan, Ji Zhicheng. Fault Diagnosis for Bearings of Unbalanced Data Based on Feature Generation[J]. Journal of System Simulation, 2020, 32(12): 2438-2448.

基于特征生成的轴承不平衡数据故障诊断

樊名鲁, 王艳, 纪志成

(江南大学物联网技术应用教育部工程研究中心, 江苏 无锡 214122)

摘要: 针对实际生产中难以获得足量的故障样本数据导致训练中样本不平衡、样本不足等问题, 提出了一种基于特征聚类的过采样算法, 并将其与卷积神经网络相结合的滚动轴承故障诊断模型。该模型将频域信号作为模型的输入, 通过卷积神经网络进行特征提取, 再通过过采样技术生成新的特征数据实现数据的均衡化, 将新生成的特征数据和原有特征一同输入到支持向量机(Support Vector Machine, SVM)分类器中完成样本的分类, 实现滚动轴承的故障诊断。通过对比实验, 结果表明该方法可以有效解决数据不平衡的问题。

关键词: 滚动轴承; 故障诊断; 特征生成; 卷积神经网络

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-731X (2020) 12-2438-11

DOI: 10.16182/j.issn1004731x.joss.20-FZ0458

Fault Diagnosis for Bearings of Unbalanced Data Based on Feature Generation

Fan Minglu, Wang Yan, Ji Zhicheng

(Engineering Research Center of Internet of Things Technology Applications Ministry of Education, Jiangnan University, Wuxi 214122, China)

Abstract: Focus on the sample imbalance and insufficiency caused by the difficulty to obtain a sufficient number of fault samples in actual production. A model for rolling bearings by combining Convolutional Neural Networks and Synthetic Oversampling is presented. The frequency domain signals is used as the input of the model, and the features are extracted by the Convolutional Neural Network. The new features are generated by Synthetic Oversampling and the data equalization is realized. The model completes the classification by putting all of the features into the Support Vector Machine, and the fault diagnosis of the rolling bearings is carried out. The comparison experiments results show that the method can effectively solve the problem of data imbalance.

Keywords: rolling bearing; fault diagnosis; feature generation; convolutional neural network

引言

滚动轴承是机械设备中最重要的零件之一, 在机械传动过程中起着传递载荷、承担应力^[1]等作用, 由于在重负载和高转速的工况下长期运行, 因

而滚动轴承是极易发生损坏的零件, 甚至造成人员伤亡。在实际应用中, 滚动轴承多在正常状态下工作, 可以收集到的故障数据往往较少, 从而出现了数据不平衡问题^[2]。数据不平衡会造成诊断准确率的下降, 因此, 提高数据失衡问题的预测性能, 尤其是少数类数据, 具有重要意义。

基于数据驱动的故障诊断是当前研究的热点, 主要涉及特征工程和分类器算法^[3]。通常提取的特征包括时域, 频域, 以及小波分解的小波特征等^[4-5]。随着深度学习的应用^[6-8], 研究者通



收稿日期: 2020-04-06 修回日期: 2020-07-08;
基金项目: 国家自然科学基金(61973138), 国家重点研发计划(2018YFB1701903);
作者简介: 樊名鲁(1996-), 男, 山东聊城, 硕士生, 研究方向为轴承故障诊断; 王艳(1978-), 女, 江苏无锡, 博士, 教授, 博导, 研究方向为制造系统能效优化。

<http://www.china-simulation.com>

• 2438 •

过将轴承振动的时域或频域信号直接输入模型中, 通过网络模型的训练得到诊断分类的结果。文献[9]提出了一种基于 LeNet-5 的新卷积神经网络进行故障诊断。首先将一维信号转换为二维图像, 然后通过提取图像的特征, 消除手工特征的影响, 最后进行故障诊断。文献[10]提出了一种基于一维卷积自编码器的检测新方法, 该方法以非线性的方式自动、精确地进行阶段划分和特征提取, 大大减少了建立模型所需的计算量, 并有效地提高了故障检测率。文献[11]提出了一种基于 RNN-LSTM 的深度学习方​​法, 其效果明显优于基于自编码方法的分类器, 并将故障诊断的错误率降到了 3%。该方法相比传统特征提取方法, 更能发现信号深层次的特征。

在复杂的时变运行条件下, 旋转设备故障的精确仿真分析是困难的。同时, 在实际操作中很难获得足够的故障样本数据而造成样本不平衡的问题。因此, 在样本不平衡下的故障诊断引起了研究人员的注意。文献[12]设计了一种振动成像方法, 该方法可以合并来自各种规模的系统中的数据。文献[13]使用残差种群相关分析提取时域特征变速箱, 然后使用 SVM 识别故障类别, 在具有噪声干扰的小样本数据的情况下, 它具有良好的故障识别能力。文献[14]通过加权少数过采样生成新的信号数据, 并通过决策树识别学习的特征, 诊断轴承是否发生故障。文献[15]首先通过主成分分析(PCA)对特征数据进行降维, 然后进行过采样合成特征, 最后通过随机森林进行故障诊断。

传统滚动轴承故障诊断是在数据集均衡的条件下进行的, 而在实际生产中需要考虑故障样本难收集的问题, 从而轴承数据是不均衡的。SVM 作为常用分类器, 具有避免维数灾难、适用小样本等优点, 被广泛用于故障诊断中。当数据集不平衡时, 经过 SVM 训练后的分类超平面会向故障样本方向偏移, 进而造成诊断准确率的下降, 甚至诊断错误。因此在不均衡数据集下, 如何提高 SVM 的故障诊断准确率是个需要考虑和解决的问题。

综合上述背景, 本文提出了一种基于特征聚类的过采样算法(Synthetic Minority Oversampling Based on Cluster, CSMO), 并将其与卷积神经网络(Convolutional Neural Network, CNN)^[16]进行结合, 将经过傅里叶变换后得到的频域信号作为模型的输入, 通过卷积神经网络完成特征的提取, 再通过基于特征聚类的过采样算法生成新的特征数据, 将新生成的特征数据和原有特征一同输入到 SVM 分类器中, 从而实现不平衡样本的轴承故障诊断。

1 问题描述

1.1 不平衡数据的轴承故障诊断问题

经过探究滚动轴承的失效机理, 我们可以清楚其失效类型主要包含点蚀, 磨损, 胶粘和塑性变形。并且, 产生故障的位置主要是在轴承的滚动元件, 内圈和外圈中。点蚀一般是比较常见和较早出现的故障, 而且点蚀的大小会随着时间而逐渐增加。图 1 为轴承点蚀发展过程的一个描述。根据点蚀的大小及位置, 我们将对其进行故障的诊断。其中, 轴承有正常(N)、内圈故障(IR)、外圈故障(OR)和滚动体故障(RB)4 种状态, 图 2 展示了每种状态下损伤尺寸为 0.007(in)下的振动信号和正常的振动信号。

另外, 在故障诊断时, 样本不平衡时会造成模型诊断能力的下降。假如有两种样本分别为正样和负样本, 当负样本占比越来越大, 模型将更多原属于正样本类的样本判别为负样本, 从而造成正样本识别错误, 诊断精度下降。当样本量少时, 模型无法深层次地发现信号中相似的特征, 对于新加入的样本, 提取出的特征区分度不高造成分类困难。

1.2 滤波级 CNN

卷积神经网络是一种结构特殊的前馈神经网络, 一般包含滤波级和分类级^[17]。本文所提模型主要使用的是滤波级, 用于从输入层提取特征, 一般模型的滤波级包含 3 层, 卷积层、激活层和池化

层。本模型在卷积层与激活层之间加入批量归一化层,可以减少内部协变量变化,提高模型的训练效率。除了滤波级,全连接层将滤波级提取出的特征进行分类,对网络参数进行训练,从而保证特征的有效性。

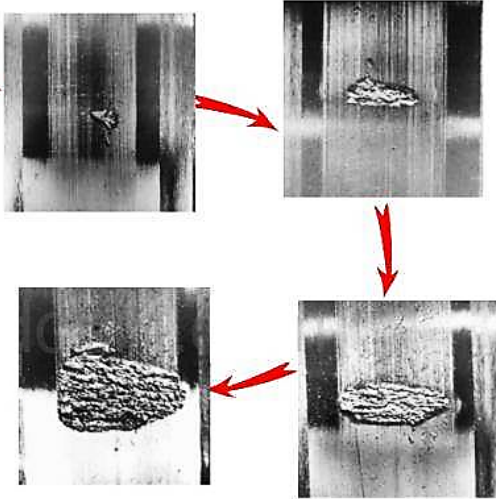


图 1 点蚀发展过程描述

Fig. 1 Pitting development process

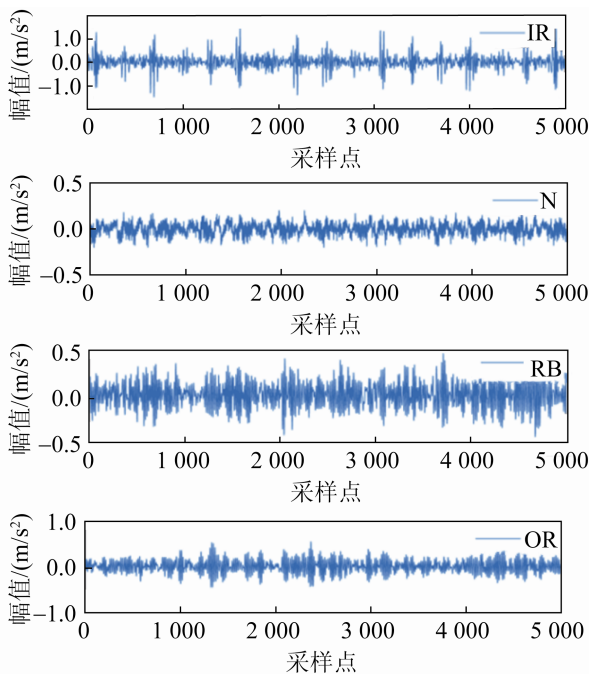


图 2 不同状态轴承的振动信号

Fig. 2 Vibration signals of bearings in different states

通过使用多个卷积核(Convolutional Kernel)对输入信号进行卷积运算并得到相应的特征,每一个

卷积核提取到一类特征。由于卷积层具有局部连接、权值共享和多卷积核等特点,卷积层可以在参数较少的情况下提取到更加丰富的特征,避免了过拟合,提高了运算效率。卷积层的输出如式(1):

$$g(i) = \sum_{x=1}^p a_x \times w_x^i + b^i, i=1,2,\dots,q \quad (1)$$

式中: i 为第 i 个卷积核; $g(i)$ 为第 i 个卷积核运算得到的特征; a 为输入数据; b 为卷积核的偏置; x 为一维输入的维度。

批量归一化(Batch Normalization, BN)层可以减少内部协变量转移,加速网络的训练过程。通常 BN 层添加在卷积层之后,激活层之前。其首要操作步骤是将卷积层的输入减去所在训练集的均值,再除以其标准差,类似于将数值进行标准化。其运算式为:

$$\hat{y}^{l(i,j)} = \frac{y^{l(i,j)} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (2)$$

式中: i 为第 i 个卷积核; $y^{l(i,j)}$ 、 $\hat{y}^{l(i,j)}$ 为标准化前后的第 j 个数值; μ_B 、 σ_B 分别为训练集的均值、标准差; ϵ 为保持数值稳定的常数项。

在进行上述操作后,会使输入值映射到一个较窄的范围。为了增强表达,将标准化得到的值进行适量的缩放,公式为:

$$z^{l(i,j)} = \gamma^{l(i)} \hat{y}^{l(i,j)} + \beta^{l(i)} \quad (3)$$

式中: $\gamma^{l(i)}$ 为缩放量; $\beta^{l(i)}$ 为偏置量; $z^{l(i,j)}$ 为经缩放操作后的值。

将批量归一化后的数据进行激活操作,它使网络获得信号的非线性表达,从而使学习的特征更加可分。修正线性单元(Rectified Linear Units, ReLU)以加快 CNN 收敛的特点被广泛使用,在反向传播训练时,ReLU 使得浅层中的权重更易训练。ReLU 的公式描述如下:

$$a^{l(i,j)} = f(z^{l(i,j)}) = \max\{0, z^{l(i,j)}\} \quad (4)$$

式中: $z^{l(i,j)}$ 为 BN 操作的输出值; $a^{l(i,j)}$ 为激活后的输出值。

通过池化层进行降采样操作,减少了网络参数的数量,并且突出了提取的特征。最大池化层是最

基本的池化层, 通过局部最大化, 减少了参数并获得位置不变的特征。最大池化描述为:

$$p_{\max}^{l(i,t)} = \max_{(j-1)w < t < jw} \{a^{l(i,t)}\}, j=1,2,\dots,q \quad (5)$$

式中: $p_{\max}^{l(i,t)}$ 为第 l 层的最大池化神经元输出; $a^{l(i,t)}$ 为第 l 层中; 第 i 个特征图的第 t 个神经元; w 为卷积核的宽度; j 为第 j 个池化核。

1.3 基于聚类的特征生成算法

我们提出了一种基于特征聚类的过采样算法 (Synthetic Minority Oversampling Based on Cluster, CSMO) 该算法首先通过模糊 C 均值聚类 (Fuzzy C-means, FCM) 对故障样本进行聚类分析, 然后通过聚类结果, 对特征样本中相似度较差的点进行识别、去除, 对于剩余的特征样本, 采用合成少数类过采样^[18] (Synthetic Minority Oversampling) 合成新的特征样本, 以使得数据集均衡。

FCM 算法本质是对目标函数的优化, 经过迭代寻找最优中心矢量, 使各样本与其距离平方和最小, 通过求极值可以确定最优模糊矩阵和聚类中心。该算法的流程如下:

对于一个 N 个样本的数据集 $X = \{x_1, x_2, \dots, x_j, \dots, x_N\}$, 并且每一样本均具有 M 个特征, 即 $x_j = (x_{j1}, x_{j2}, \dots, x_{jM})$ 。首先进行初始化, 将数据集 X 划分为 $C(2 \leq C \leq N)$ 个模糊子集, 其中 C 是聚类数目, 并将得到的初始聚类结果用一个聚类中心矢量 V 和隶属度矩阵 U 表示, 即

$$V = \{v_1, v_2, \dots, v_i, \dots, v_C\} \quad (6)$$

式中: $v_i = (v_{i1}, v_{i2}, \dots, v_{ik}, \dots, v_{iM})$ 为第 i 类的聚类中心, v_{ik} 为它的第 k 维特征;

式(7)中, u_{ij} 为第 j 个样本相对于第 i 类的隶属度, 且 u_{ij} 满足式(8)

$$U = \begin{bmatrix} u_{11} & \cdots & u_{1j} & \cdots & u_{1N} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ u_{i1} & \cdots & u_{ij} & \cdots & u_{iN} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ u_{C1} & \cdots & u_{Cj} & \cdots & u_{CN} \end{bmatrix} \quad (7)$$

$$\begin{cases} \sum_{i=1}^C u_{ij} = 1, & j=1,2,\dots,N \\ u_{ij} \in [0,1], & i=1,2,\dots,C; j=1,2,\dots,N \\ \sum_{j=1}^N u_{ij} \in [0,N], & i=1,2,\dots,C \end{cases} \quad (8)$$

接下来需要从其分类空间中寻找一个最优结果。因此, 定义 FCM 算法的目标函数为:

$$J(U, V) = \sum_{j=1}^N \sum_{i=1}^C (u_{ij})^m (d_{ij})^2 \quad (9)$$

式中: m 为影响隶属度矩阵模糊化程度的权重指数, 且 $m \in [0, \infty]$; d_{ij} 为样本到聚类中心的距离, 可以表示为:

$$d_{ij} = \|x_j - v_i\|_A = (x_j - v_i)^T A (x_j - v_i) \quad (10)$$

式中: A 为 M 阶的正定对称矩阵, 当 A 取单位矩阵时, 式(10)即为欧氏距离。

为了在 u_{ij} 满足(8)的条件下, 求式(9)的极小值, 构造拉格朗日函数, 即

$$L(U, V, \lambda) = \sum_{j=1}^N \sum_{i=1}^C (u_{ij})^m (d_{ij})^2 + \sum_{j=1}^N \lambda_j \left(\sum_{i=1}^C u_{ij} - 1 \right) \quad (11)$$

式中: λ_j 为拉格朗日乘子。分别求 $L(U, V, \lambda)$ 对 u_{ij} , v_i 和 λ_j 的偏导数, 经计算, 可解得当式(11)取极小值时的模糊隶属度矩阵 U 和聚类中心矢量 V 分别为:

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}}}, i=1,2,\dots,C; j=1,2,\dots,N \quad (12)$$

$$v_i = \frac{\sum_{j=1}^N (u_{ij})^m x_j}{\sum_{j=1}^N (u_{ij})^m}, i=1,2,\dots,C \quad (13)$$

SMO 是基于随机过采样的一种改进算法, 对于少数类样本, 随机过采样采用复制的方式处理, 这样容易使模型过拟合。SMO 则利用 K 近邻和线性插值, 在两个少数类样本间按照一定的规则插入新的样本。算法具体实现过程如下:

(1) 对于少数类中每一个样本 x , 计算其欧氏距离, 得到它到其他所有少数类样本的距离, 得到

其 K 近邻。

(2) 根据样本不均衡程度设置一个采样倍率 N ，对每一个少数类样本 x ，从其 K 近邻中随机选择若干样本，假设其中一个选中的近邻为 x_n 。

(3) 对于每一个随机选择的近邻 x_n ，按照如下的公式(14)构建新的样本 X_{new} ：

$$X_{\text{new}} = x + \text{random}(0,1) \times (x_n - x) \quad (14)$$

(4) 把人工合成的新样本与原始数据集组成一个新的数据集，然后利用分类器对其进行分类。

2 基于 CNN 和 CSMO 的不均衡轴承故障诊断模型

2.1 故障诊断流程

对于不均衡样本的轴承故障诊断模型主要包括 CNN 特征提取模型与 CSMO 集成 SVM 分类器两部分。通过 CNN 实现了特征的自适应提取，并且能发现信号中深层次的特征。通过 CSMO 集成 SVM 分类器，合成了新的特征样本，使样本更加均衡，并且 SVM 相比其它分类器，具有更好的诊断准确率。该诊断模型的流程为：将经过傅里叶变换后得到的频域信号输入到 CNN 中进行训练，根据 Softmax 分类结果对 CNN 网络进行参数调整，直到结果稳定，得到训练集中提取的特征；通过 CSMO 对不均衡特征样本进行过采样，输入到 SVM 分类器中进行训练；利用训练好的 CNN 特征提取模型与 SVM 分类器对测试集进行故障诊断，模型流程图如图 3 所示。

传统特征提取需要人为确定特征类别，并且人工提取的特征不能完全反映信号，通过深度学习能够深层次地挖掘数据间的联系，CNN 作为常用的深度学习模型可以满足要求。本文通过过采样的方法进行特征的生成，相对于只是在算法上进行改进提高准确率，该方法提供了另外一种思路，并且进行了实验论证。鉴于对于小样本的问题，SVM 分类器一直有较好的表现，于是对 CNN 模型进行改进，使用 SVM 分类器代替 Softmax 函数，有效地提高了分类结果的准确性。

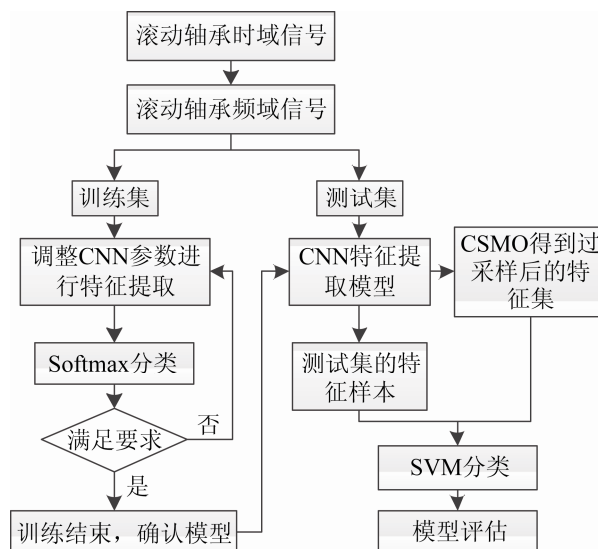


图 3 基于 CNN 和 CSMO 的诊断流程

Fig. 3 Diagnostic process based on CNN and CSMO

2.2 基于 CNN 的特征提取

特征提取是该诊断模型中重要的阶段，该模型使用 CNN 进行特征提取，其中 CNN 的具体结构如图 4 所示。

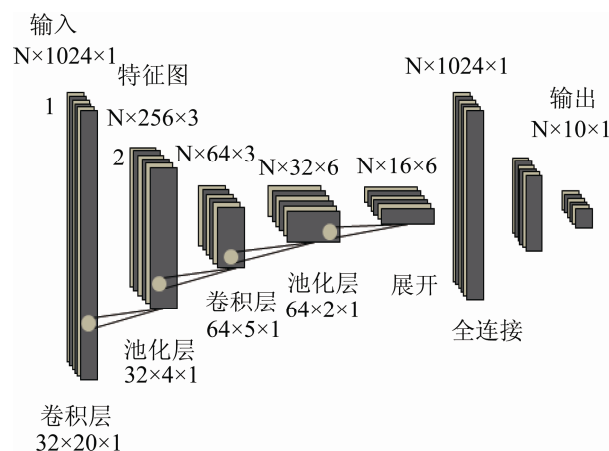


图 4 深度卷积特征提取模型结构图

Fig. 4 Structure of deep convolution feature extraction model

在本 CNN 模型中，一共进行了 2 次滤波操作，第一次滤波操作的输入是 3 维张量，其维度为 (None, 1024, 1)，None 代指输入 CNN 的样本数。输入先经过一维卷积层，该层由 32 个 20×1 的卷积核组成，步长为 4×1，再经过一层最大值池化层，池化尺寸为 4×1，步长为 4×1，则该层的输出维度

为(None, 64, 32)。第二次滤波操作的由 64 个 5×1 的卷积核组成, 步长为 2×1 , 池化尺寸为 2×1 , 步长为 2×1 , 该层的输出维度为(None, 16, 64)。每一次滤波操作的激活函数为 RELU, 同时添加了批量归一化层以减少内部协变量转移。然后经过 3 层全连接层, 神经元个数分别为 500、200 和 20。

在模型确定后, 需要对模型进行训练, 通过反向传播对网络进行权值优化。主要步骤是先求出目标函数对最后一层神经元的导数, 然后从后向前, 算出目标函数对于每层权值的导数值。为了使目标函数最小, 通过训练寻找最优的权值, 该过程如式(15)所示。

$$\theta^* = \arg \min_{\theta} L(f(x^i; \theta)) \quad (15)$$

式中: $L(\cdot)$ 为目标函数值; $f(\cdot)$ 为输出值; θ 为网络的所有参数; x^i 为网络的输入; θ^* 为网络的最优参数。

本文采用 Adam(Adaptive Moments)算法进行参数优化。Adam 通过计算梯度的一阶矩估计和二阶矩估计而为不同的参数设计独立的自适应性学习率, 动态地更新参数的学习率。Adam 对超参数的选择具有鲁棒性, 因此有助于参数优化, Adam 算法流程为:

- 1: 输入: 步长 λ , 矩估计指数衰减率 ρ_1 和 ρ_2 , 数值稳定常数 ε , 初始化一阶和二阶矩变量 $s_0=0$, $v_0=0$, 初始化时间步 $t=0$
- 2: 初始化: 初始变量: θ_0
- 3: 过程:
- 4: while 参数 θ_t 没有达到停止准则 do
- 5: $t \leftarrow t + 1$
- 6: $g_t \leftarrow L_t(f(x^i; \theta))$
- 7: $s_t \leftarrow \rho_1 s_{t-1} + (1 - \rho_1) g_t$
- 8: $v_t \leftarrow \rho_2 v_{t-1} + (1 - \rho_2) g_t^2$
- 9: $\hat{s}_t \leftarrow \frac{s_t}{1 - \rho_1^t}$
- 10: $\hat{v}_t \leftarrow \frac{v_t}{1 - \rho_2^t}$
- 11: $\theta_t \leftarrow \theta_{t-1} - \lambda \frac{\hat{s}_t}{\sqrt{\hat{v}_t} + \varepsilon}$

12: end while

13: return θ_t

2.3 特征分类

在得到特征提取模型之后, 对测试集进行特征提取, 对特征进行过采样, 得到均衡样本, 再输入到 SVM 中进行分类。SVM 是依据统计学习理论, 建立在结构风险最小化原则基础上的分类算法。其采用的最优决策函数为式(16):

$$f(x) = \text{sgn} \left(\sum_{i=1}^N a_i^* y_i K(x, x_i) + b^* \right) \quad (16)$$

式中: $\text{sgn}(\cdot)$ 为符号函数; $K(x, x_i)$ 为所选径向基核函数(Radial Basis Function, RBF), 即

$$K(x, x_i) = \exp \left(-\frac{\|x - x_i\|^2}{2\sigma^2} \right) = \exp(-g \|x - x_i\|) \quad (17)$$

式中: σ 为径向基函数的宽度, σ 越小, 径向基函数的宽度越小, 越有选择性。

RBF 仅仅有一个参数, 即 σ 。而在 SVM 中, 惩罚因子 C 的选择能够表现分类最大化间隔, 避免错分情况, 对分类结果有着较大的影响。一定区间上的参数 C 和 σ 会有较高的分类准确率, 但在多数范围内, 分类准确率却很低, 因此进行参数的选择还是很有必要的。

网格搜索法是让 C 和 σ 在一定范围划分网格, 并取值网格内所有点。对于给定的 C 和 σ 进行分类训练, 得到在此组 C 和 σ 下训练集分类准确率, 最终, 取训练集分类准确率最高的那组 C 和 σ 为模型参数。其中 C 的范围设置为 $[2^{-10}, 2^7]$, σ 的范围设置为 $[2^{-10}, 2^3]$, 步距为 0.1。

3 算例仿真及验证

3.1 数据样本及仿真环境

本实验采用从美国西储大学获得的轴承故障数据集, 选用实验装置驱动端(Drive Ending, DE)型号为 SKF-6205 的轴承, 其转速为 1 750 r/min, 信号采样频率为 48 kHz, 轴承有正常(N)、内圈故障(IR)、外圈故障(OR)和滚动体故障(RB) 4 种状态。

该实验集引入单点故障来模拟滚动轴承的点蚀故障。通过使用放电加工方法,在目标轴承上加工出不同尺寸的凹坑。分别在轴承的滚动元件,内圈和外圈上加工出断层直径为 0.007(in), 0.014(in)和 0.021(in)的坑。取不同的断层直径,表示了点蚀断层从微断层到显着断层逐渐演化的过程。根据故障不同将轴承状态划分为 10 种,即一共有 10 种不同的训练和测试集。在进行实验时,同时取 5~7 转的振动信号,即取每段样本长度为 2 048 点的信号数据输入到模型中。

3.2 CNN 模型的优化

特征提取器 CNN 的性能对 CNN-SVM 模型的最终诊断结果有很大的影响。在构建 CNN 模型的过程中,选择合适的模型超参数可以优化模型,提升训练速度和训练准确性。本小节通过实验论证了模型的建立,研究了优化器、模型结构、特征维度等指标对模型性能的影响。该部分实验以最终模型为参考模型,通过改变对应变量进行论证。

(1) 优化算法的选择

对于深度学习模型,选择合适的优化算法能够提高模型的训练速度和分类精度。在此实验中,本文测试并比较了 4 个最常用的优化算法,包括 SGD, Adadelta, RMSProp 和 Adam 优化算法。图 5 显示了前 50 次取样中 4 种优化算法的测试精度曲线,其中迭代 50 次后进行一次取样。

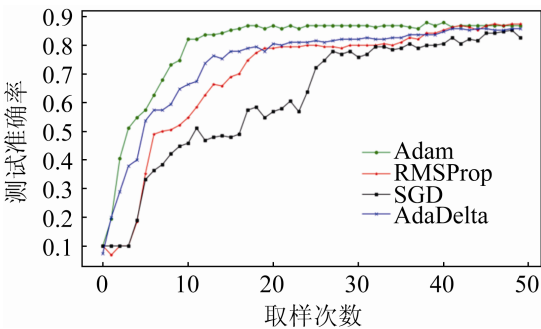


图 5 优化算法性能比较

Fig. 5 Optimal algorithm performance comparison

从图 5 可以清楚地看到 Adam 优化器的性能明

显优于其他优化器。特别是在迭代计算的前 20 个阶段,亚当优化算法的训练精度迅速提高,并且后期能够更加稳定地收敛。在训练准确性方面,4 种算法效果相近,Adam 优化器整体准确性略高,而 SGD 效果在此实验中最差。

(2) CNN 模型结构的优化

为了研究不同的 CNN 模型结构对准确性性能的影响,在此实验中,本文设计了 3 种 CNN 模型结构,包括一层 CNN,两层 CNN 和三层 CNN 结构。表 1 列出了不同 CNN 模型的超参数。不同的 CNN 模型的精度曲线如图 6 所示。

表 1 3 种 CNN 模型结构	
Tab. 1 Three kinds of CNN model structure	
模型 层数	CNN 模型参数
L=1	卷积层(32, 20×1, 4)+池化层(4×1, 4)+全连接层(500/200/20)+Softmax
L=2	CNN(L=1)+卷积层(64, 5×1, 2)+池化层(2×1, 2)+全连接层(500/200/20)+Softmax
L=3	CNN(L=2)+卷积层(64, 5×1, 2)+池化层(2×1, 2)+全连接层(500/200/20)+Softmax

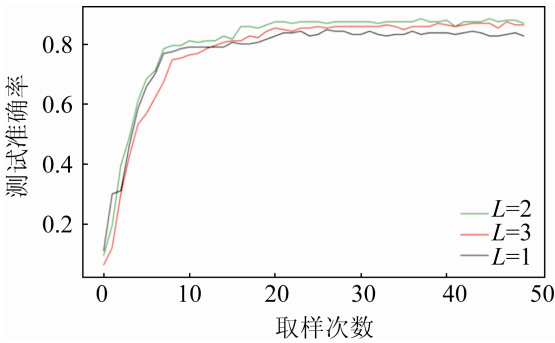


图 6 模型结构性能比较

Fig. 6 Vibration signals of bearings in different states

从图 6 可以看出,当模型是两层结构时,具有最佳的性能,一层 CNN 模型的性能最差,由于结构简单,特征提取能力不足。在该实验中,由于使用小样本数据集进行训练,因此三层 CNN 的模型比两层 CNN 模型训练速度慢。

(3) 特征维度的确定

特征维度的大小影响着模型的性能。若特征维

度太大, 将会有部分冗余特征存在, 并且影响 CSMO 生成样本的质量; 若特征维度过小, 将无法反映信号的全部信息, 影响分类的准确率。在本实验中, 通过改变倒数第二层全连接层的输出维度来改变特征维度, 探究其对诊断的影响。该层输出神经元个数分别设置为 10、15、20、25 和 30。实验数据的设置和 3.2 节相同, 其中不平衡率设置为 10:3。实验结果如表 2 所示。

表 2 维度对比结果
Tab. 2 Result of dimension comparison

特征维度	10	15	20	25	30
CNN	80.88	84.04	83.16	82.98	83.33
CNN-CSMO	87.54	88.25	91.40	89.47	89.65

由表 2 可以看出, 对于原始模型, 当特征维度为 15 时诊断准确率最高, 而当特征维度为 20 时, 所提出的模型的诊断准确率最高, 所以模型倒数第二层的神经元个数设置为 20。

3.3 不平衡度对诊断的影响

本文解决的是不平衡样本下的故障诊断问题, 为了证明算法的有效性, 本节对不同比例下的样本进行了诊断。在所有样本中, 有一部分类别的样本数量是足够的, 在本实验中, 样本足够的类别(大类)取 50 个样本, 而样本较少的类别(小类)的样本数量根据不同比例发生变化, 假设不平衡率为 N, 具体的数据划分如表 3 所示。

表 3 数据集信息
Tab. 3 Information of data set

故障类型	标签	训练/个	测试/个
正常	9	50	20
内圈故障(0.021)	8	50	20
外圈故障(0.021)	7	50	20
滚动体故障(0.021)	6	50	20
内圈故障(0.014)	5	50	20
外圈故障(0.014)	4	50/N	20
滚动体故障(0.014)	3	50/N	20
内圈故障(0.007)	2	50/N	20
外圈故障(0.007)	1	50/N	20
滚动体故障(0.007)	0	50/N	20

在本节的实验中, 数据不平衡率分别设置为 10:1、5:1、10:3、2:1 和 1:1。为了减少随机性的影响, 在每种情况下均重复进行 5 次实验, 并将平均值作为最终故障诊断的结果。为了模拟实际情况下的噪声干扰, 将信号加上 SNR=-6 db 的高斯噪声。其中下采样技术和本文所提到的过采样技术是对抗不平衡样本的常用方法。因此, 我们将本文所提方法 CNN-CSMO 的诊断准确性分别与 CNN(无采样), CNN(下采样), 进行了比较。表 4 显示了不同方法、不同比例下的故障诊断准确率。

表 4 不平衡率对比结果
Tab. 4 Result of imbalance comparison

诊断方法	10:1	5:1	10:3	2:1	1:1
CNN-CSMOTE	88.07	87.37	91.40	90.53	无
CNN(无采样)	80.70	81.75	83.16	84.74	89.12
CNN(下采样)	76.49	78.95	81.05	82.63	无

从实验结果可以看出, 在该诊断模型中, 下采样并不是解决不平衡问题的好方法, 样本的减少降低了模型的诊断准确率。而本文提出的方法有较好的诊断性能, 随着不平衡率的下降, 诊断准确率也得到了提高。由结果可以看出, 当不平衡比为 10:1 时, 即小类样本数最少时, 诊断准确率提高了 8 个百分点, 几乎达到了均衡样本的诊断准确率。另外, 通过本文方法, 可以达到比均衡样本更高的诊断准确率, 考虑是分类器的改变起到了作用, 关于这一点, 在后面还将进一步验证。

3.4 与其他方法的比较

(1) 不同分类器的对比

本节对不同分类器进行了实验, 除了 SVM, 还有后向传播神经网络(BPNN)、随机森林(RF)、自适应增强学习(Adaboost)等常用分类器。对不同分类器在噪声环境下的诊断准确率对比, 发现 SVM 是该模型中最有效的分类算法, 实验结果如表 5 所示。这也是所提模型在不均衡条件下比原模型准确率高的原因。

表 5 分类器对比结果
Tab. 5 Comparison result of classifier

分类方法	-10 dB	-8 dB	-6 dB	-4 dB	-2 dB
BPNN	82.81	84.56	88.95	90.35	92.28
SVM	85.09	88.95	91.93	92.28	94.56
RF	82.63	85.61	90.35	91.58	93.68
Adaboost	75.62	78.21	84.56	88.95	91.58

(2) 不同特征提取方法的对比

对比实验选取基于传统特征提取的一类诊断方法和自适应特征提取的深度学习诊断方法。前者中提取了 14 种时域特征、4 种频域特征和通过集成经验模态分解得到的 5 个时频域特征进行故障分类, 分类器采用 SVM 分类器; 后者中选取了特征提取能力较强的堆叠降噪自编码网络(SDAE, 结构 1024-256-128-10)。为了验证本文模型的抗噪性能, 本节实验将在 SNR 为-10 dB、-6 dB、-2 dB 的环境下进行。

实验结果如表 6 所示。由结果可以看出, 所提出的模型有较好的抗噪声性能, 在噪声较大时能够保持较高的诊断准确率。当噪声较小时, SDAE 模型展现出了较好的特征提取能力, 但本文模型效果也不错。

表 6 模型对比结果
Tab. 6 Result of model comparison

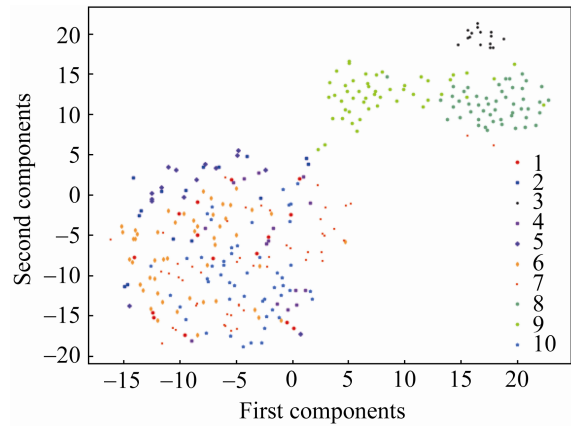
诊断方法	-10 dB	-6 dB	-2 dB
SVM	65.2	80.53	88.42
SVM-SMOTE	66.32	81.05	90.53
SDAE	64.74	90.53	94.74
SDAE-SMOTE	67.9	92.11	95.26
CNN	65.09	83.16	86.84
CNN-SMOTE	85.09	91.4	93.68

(3) 诊断结果分析

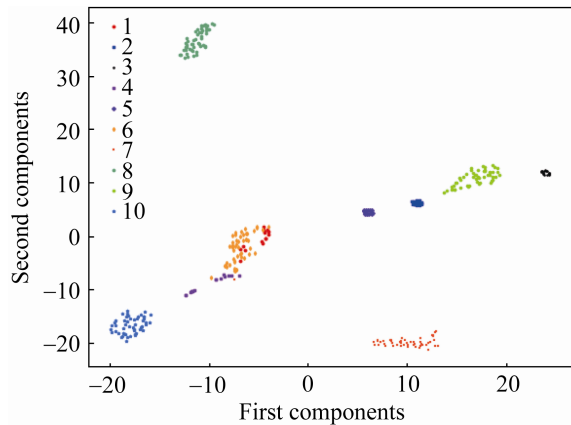
为了更好地理解不同模型的特征提取能力, 本实验采用主成分分析(PCA)对人工提取的特征、通过卷积神经网络自适应提取的特征以及经过过采样后的特征进行分析。其中, 实验是在信噪比为-6 dB 的情况下进行的, 横纵坐标分别代表前两维主分量, 分析结果如图 7 所示。

对比图 7(a)与图 7(b), 卷积神经网络在自适应

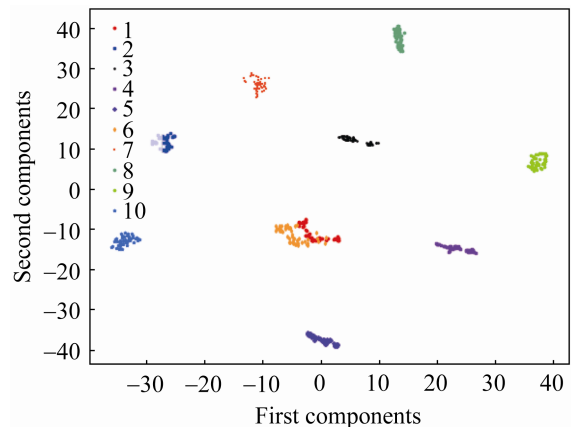
特征提取上取得了很好的效果, 提取效果明显优于人工提取的特征。对比图 7(b)与图 7(c), 会发现过采样后的特征区分度得到了进一步提高, 有助于分类准确度的提升。



(a) 人工提取的特征



(b) CNN 提取的特征



(c) CNN-SMOTE 提取的特征

图 7 诊断结果分析

Fig. 7 Analysis of Diagnosis

4 结论

实际生产中存在难以获得足量的故障样本数据导致训练中样本不平衡等问题。针对此类问题: 提出了一种基于特征聚类的过采样算法, 并将其与卷积神经网络相结合的滚动轴承故障诊断方法, 该方法通过卷积神经网络实现了特征的自适应提取, 并通过过采样算法进行特征生成实现数据的均衡化, 最后将新生成的特征数据和原有特征一同输入到 SVM 分类器中完成样本的分类, 实验证明该模型能够在样本不平衡时提高故障的诊断准确率, 对小样本有较好的效果。通过与基于传统特征提取的诊断模型和自适应特征提取的深度学习诊断模型的对比实验, 表明模型在具体故障类别诊断中具备准确率高的优势, 并且模型具有更好的抗噪声能力。

通过实验发现, 在噪声、不平衡样本环境下所提出模型有一定的效果, 但若故障诊断率达到 95% 以上时, 该模型很难提高故障诊断率。所以, 如何提高模型的特征提取能力, 提高假样本的真实性, 使模型能够提取更多有用的特征将是下一步需要解决的问题。

参考文献:

- [1] 阳建宏, 黎敏, 丁福焰. 滚动轴承诊断现场实用技术[M]. 北京: 机械工业出版社, 2015.
Yang Jianhong, Li Min, Ding Fuyan. Field Practical Technology of Rolling Bearing Diagnosis[M]. Beijing: China Machine Press, 2015.
- [2] Razavi-Far R, Farajzadeh-Zanjani M, Saif M. An Integrated Class-imbalanced Learning Scheme for Diagnosing Bearing Defects in Induction Motors[J]. IEEE Transactions on Industrial Informatics, 2017, 13(6): 2758-2769.
- [3] 文成林, 吕菲亚, 包哲静, 等. 基于数据驱动的微小故障诊断方法综述[J]. 自动化学报, 2016, 42(9): 1285-1299.
Wen Chenglin, Lü Feiya, Bao Zhejing, et al. A Review of Data Driven-based Incipient Fault Diagnosis[J]. Acta Automatica Sinica, 2016, 42(9): 1285-1299.
- [4] 徐继亚, 王艳, 纪志成. 基于鲸鱼算法优化 WKELM 的滚动轴承故障诊断[J]. 系统仿真学报, 2017, 29(9): 2189-2197.
Xu Jiya, Wang Yan, Ji Zhicheng. Fault Diagnosis Method of Rolling Bearing Based on WKELM Optimized by Whale Optimization Algorithm[J]. Journal of System Simulation, 2017, 29(9): 2189-2197.
- [5] 王田田, 王艳, 纪志成. 基于改进极限学习机的滚动轴承故障诊断[J]. 系统仿真学报, 2018, 30(11): 4413-4420.
Wang Tiantian, Wang Yan, Ji Zhicheng. Fault Diagnosis of Rolling Bearing Based on Improved Extreme Learning Machine[J]. Journal of Vibration and Shock, 2018, 30(11): 4413-4420.
- [6] 任浩, 屈剑锋, 柴毅. 深度学习在故障诊断领域中的研究现状与挑战[J]. 控制与决策, 2017, 32(8): 1345-1358.
Ren Hao, Qu Jinfeng, Chai Yi. Deep Learning for Fault Diagnosis: The State of the Art and Challenge[J]. Control and Decision, 2017, 32(8): 1345-1358.
- [7] Wen L, Li X, Gao L, et al. A New Convolutional Neural Network Based Data Driven Fault Diagnosis Method[J]. IEEE Transactions on Industrial Electronics (S 1557-9948), 2018, 65(7): 5990-5998.
- [8] Sun J, Yan C, Wen J. Intelligent Bearing Fault Diagnosis Method Combining Compressed Data Acquisition and Deep Learning[J]. IEEE Transactions on Instrumentation and Measurement (S1557-9662), 2018, 67(1): 185-195.
- [9] Guo L, Lei Y, Xing S, et al. Deep Convolutional Transfer Learning Network: A New Method for Intelligent Fault Diagnosis of Machines with Unlabeled Data[J]. IEEE Transactions on Industrial Electronics (S 1557-9948), 2019, 66(9): 7316-7325.
- [10] 王硕, 王培良. 基于一维卷积自编码器—高斯混合模型的间歇过程故障检测[J]. 信息与控制, 2019, 48(3): 285-292.
Wang Shuo, Wang Peiliang. Fault Detection Method for a Batch Process Based on a One dimensional Convolution Autoencoder and Gaussian Mixture Model[J]. Information and Control, 2019, 48(3): 285-292.
- [11] 曲星宇, 曾鹏, 徐承成, 等. 基于 DropOut 降噪自编码的磨矿系统故障诊断[J]. 控制与决策, 2018, 33(9): 1662-1666.
Qu Xingyu, Zeng Peng, Xu Chengcheng, et al. DropOut Denoising Autoencoder-based Fault Diagnosis for Grinding System[J]. Control and Decision, 2018, 33(9): 1662-1666.

- [12] Oh H, Jung J H, Jeon B C, et al. Scalable and Unsupervised Feature Engineering using Vibration-imaging and Deep Learning for Rotor System Diagnosis[J]. IEEE Transactions on Industrial Electronics (S1557-9948), 2018, 65(4): 3539-3549.
- [13] Jiao W, Lin S. Overall-improved Fault Diagnosis Approach Based on Support Vector Machine[J]. Chinese Journal of Scientific Instrument (S0254-3087), 2015, 36(8): 1861-1870.
- [14] Zhang Y Y, Li X Y, Gao L, et al. Imbalanced Data Fault Diagnosis of Rotating Machinery using Synthetic Oversampling and Feature Learning[J]. Journal of Manufacturing Systems (S0278-6125), 2018, 48(4): 34-50.
- [15] Hang Q, Yang J, Xing L. Diagnosis of Rolling Bearing Based on Classification for High Dimensional Unbalanced Data[J]. IEEE Access (S2169-3536), 2019, 7(9): 79159-79172.
- [16] Hassani H, Zarei J, Arefi M M, et al. Zslices-based General Type-2 Fuzzy Fusion of Support Vector Machines with Application to Bearing Fault Detection[J]. IEEE Transactions on Industrial Electronics (S1557-9948), 2017, 64(9): 7210-7217.
- [17] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature (S1476-4687), 2015, 521(7553): 436-444.
- [18] Chawla N V, Bowyer K W, Hall L O, et al. SMOTE: Synthetic Minority Over-sampling Technique[J]. Journal of Artificial Intelligence Research (S1076-9757), 2002, 16(1): 321-357.