

DOI: 10.19650/j.cnki.cjsi.J1905073

基于深度学习的电机轴承微小故障智能诊断方法^{*}

宫文峰^{1,2}, 陈辉¹, 张美玲², 张泽辉¹

(1. 武汉理工大学 高性能舰船技术教育部重点实验室 武汉 430063; 2. 桂林电子科技大学北海校区 北海 536000)

摘要:运用深度学习技术对滚动轴承微小故障发生的位置、类别和严重程度进行精准自动的辨识是当前故障诊断领域研究的热点。传统的故障诊断方法过度依赖于工程师凭经验进行手工特征提取,难以有效提取微小故障特征。提出了一种改进的CNNs-SVM的新方法用于电机轴承的故障快速智能诊断,该方法采用 1×1 的过渡卷积层与全局均值池化层的组合代替传统CNN的全连接网络层结构,有效减少CNN的训练参数量,在测试阶段采用支持向量机代替Softmax分类器进一步提升诊断准确率。最后将提出的方法用于电机支撑滚珠轴承的故障实验数据并与多种算法对比验证。结果表明,改进CNNs-SVM算法的故障识别准确率高达99.86%,同时在不同负载下具有良好的迁移泛化能力,具备实际工程应用的可行性。其诊断准确率和测试时间明显优于其他智能算法。

关键词:故障诊断;卷积神经网络;支持向量机;深度学习;全局均值池化

中图分类号: TH165+.3 TP277 TP306+.3 文献标识码: A 国家标准学科分类代码: 520.20

Intelligent diagnosis method for incipient fault of motor bearing based on deep learning

Gong Wenfeng^{1,2}, Chen Hui¹, Zhang Meiling², Zhang Zehui¹

(1. Key Laboratory of High Performance Ship Technology of Ministry of Education, Wuhan University of Technology, Wuhan 430063, China; 2. Beihai Campus, Guilin University of Electronic and Technology, Beihai 536000, China)

Abstract: Using deep learning technique to automatically and accurately identify the incipient fault of rolling bearing, especially the fault position, classification and severity degree, is a research hotspot in current fault diagnosis field. The traditional fault diagnosis method excessively relies on the manual feature extraction by the engineers with prior knowledge, which is difficult to effectively extract incipient fault features. In this paper, a novel improved CNNs-SVM method is proposed and used for the rapid intelligent fault diagnosis of motor rolling bearing. This method adopts the combination of 1×1 transitional convolution layer and global average pooling layer to replace the fully connected network layer structure of traditional CNN, which effectively reduces the number of training parameters of CNN. In test stage, the method uses SVM to replace the Softmax classifier, which further improves the diagnosis accuracy. The proposed method was applied to the fault experiment data of the motor support rolling bearing, and the method was compared and verified with traditional intelligent diagnosis methods. The results show that the accuracy of fault identification of the improved CNNs-SVM algorithm reaches up to 99.86%, and the proposed method has good migration generalization ability under different load conditions and possesses the feasibility for practical engineering application. The fault diagnosis accuracy and test time of the method is obviously better than other intelligent algorithms.

Keywords: fault diagnosis; convolutional neural network; support vector machine; deep learning; global average pooling

0 引言

近年来,我国在国产航空母舰、新型战略核潜艇、高

性能电力推进船等为代表的海洋工程装备研制方面取得了一系列突破性进展^[1-2]。随着高性能舰艇朝向高速化、智能化和自动化方向快速发展,建立高效智能的故障诊断和健康监测系统是保证重大装备正常可靠运行的重要

收稿日期: 2019-04-29 Received Date: 2019-04-29

^{*} 基金项目: 国家自然科学基金面上项目(51579200)、中央高校基本科研业务经费资助武汉理工大学优秀博士学位论文培育项目(2019-YB-023)、广西高校中青年教师科研基础能力提升项目(2019KY0216)、国家留学基金委博士联合培养项目资助

措施^[3]。电机是高性能电力推进舰船最重要的末端执行部件,经研究统计表明,滚动轴承作为推进电机中的核心支撑回转部件,受冲击交变载荷、热疲劳和机械磨损等因素影响,是机械故障最易发的关键部件^[4-5],其健康状态直接影响装备的总体性能。因此,针对滚动轴承的故障诊断方法一直是当前领域内学者研究的热点^[4-6]。

目前主流的故障诊断方法主要分为基于解析模型^[7]和基于数据驱动^[8-9]两种方法。基于模型的方法通常需要建立高精度的数学模型,对于复杂的多变量系统建模代价高,且建立的模型在解决同类问题上普适性较差^[9],在实际应用中具有局限性。如今随着先进传感器和智能检测技术成群组的应用于机械健康监测中,反映装备全生命周期健康状态的振动、噪声和电力参数等监测“大数据”被轻易获取,因此基于数据驱动的故障诊断方法在工业界得到广泛应用^[6,10]。近10年来,随着机器学习研究的不断兴起,基于人工智能的数据驱动故障诊断算法成为故障诊断领域研究的热点^[9],比较有代表性的支持向量机(support vector machine, SVM)、BP神经网络(BPNN)和K近邻法(K-nearest neighbor, KNN)已被广泛应用于故障诊断,并取得了较好的效果^[11-13]。经过多年的应用验证,以上3种算法因其浅层的网络结构导致特征提取能力不足,从而难以挖掘和提取隐藏于故障数据中更深层次的微小特征^[5],限制了诊断准确率的进一步提升。

2006年,Hinton等^[14]提出深度学习理论,它利用深层神经网络对输入样本数据逐层贪婪学习并自动提取代表性特征,目前已成功应用于图像识别、语音识别和自然语言处理等领域^[15],近5年开始受到故障诊断领域学者的关注^[16-17]。卷积神经网络(convolutional neural network, CNN)^[10]是深度学习重要的分支之一,目前主要应用于图像识别。CNN强大的特征提取能力具备识别微小故障特征的潜力,克服了传统智能诊断方法中的固有缺陷^[5],近5年来,部分学者已将CNN应用于故障诊断领域,文献[10]和[18]分别将CNN用于滚动轴承和异步电机的故障诊断中,两者虽然都用到了CNN算法,但是仍然需要借助传统的FFT方法等对原始数据进行特征提取,从而未充分利用CNN强大的特征提取能力,且限制了诊断效果的进一步提升。文献[19]提出了一种无手工特征提取的CNN方法用于滚动轴承故障诊断,但是该CNN模型的全连接层的训练参数量过大,导致训练和测试时间较长,不利于微小故障的快速诊断。

针对以上不足,本文提出了一种基于改进CNNs-SVM的新方法用于电机支撑轴承的微小故障智能诊断。本文方法是基于改进的CNNs作为特征提取器和采用SVM作为故障分类器相结合,通过混淆矩阵对故障误判的位置和数量进行量化。其次,通过设计 1×1 的过渡卷积层与全局均值池化层组合代替传统CNN中2~3层的

全连接网络结构,极大减少了训练参数量和计算时间,并运用多种深度学习训练技巧提高了算法的鲁棒性和通用性,使得提出的方法更适用于故障的快速诊断和监测。

1 CNNs-SVM 智能故障诊断算法

1.1 传统 CNN 算法模型

CNN是一种专门用于处理具有类网状结构数据的前馈式神经网络模型^[20],具有强大的特征提取能力,可以将输入层原始数据逐层转化为更易于识别的特征。CNN通过构建多个能够提取数据特征的滤波器(卷积核),并运用这些滤波器对输入数据逐层提取隐藏在数据中的代表特征,同时结合稀疏式连接和参数权共享机制在时间和空间上降采样精减数据维度,极大减少了训练参数量^[21]。CNN模型的基本架构如图1所示。

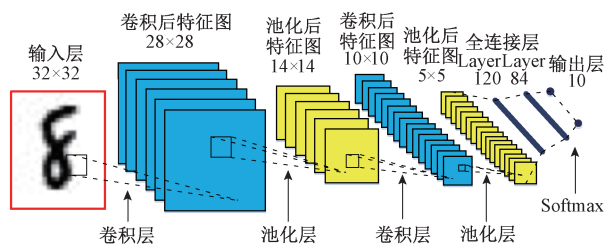


图1 卷积神经网络的基本结构

Fig.1 The basic structure of convolutional neural network

CNN采用反向传播算法将误差逐层反向传递,并使用梯度下降法调整各层之间的参数^[19]。由图1可知,CNN一般主要由卷积层、激活层、池化层、全连接层和Softmax分类层组成^[18]。通过交替堆叠多个卷积、激活和池化组合即可构建深层CNN网络^[4],其前向计算过程如式(1)所示:

$$f(X) = f_N(\cdots f_3(f_2(f_1(X, \theta^{(1)}), \theta^{(2)}), \theta^{(3)}), \cdots), \theta^{(N)}) \quad (1)$$

式中: X 表示原始的输入数据; $\theta^{(N)}$ 为第 N 层所学习到的训练参数值,如权重和偏置; f_N 为第 N 层所执行的运算操作; $f(X)$ 为原始数据 X 经上述卷积、激活和池化操作得到的输出特征图。

在图1所示的传统CNN网络中,判别输出层通常由一个全连接层与Softmax分类器组合构成^[18],其全连接层通常为2~3层的全连接网络,然而,全连接层所带入的参数数量几乎占据了CNN模型总参数量的80%~90%^[22-23],既占用过多的计算资源又容易引起过拟合,使CNN减少训练参数量的优势又变得微弱不明显。另外,Softmax分类器实质上是对最终的分类结果做一次符合概率分布的归一化操作^[18],在多元分类性能上尚不如SVM的功能强大。然而SVM的不足之处在于其深层特征提

取和数据挖掘能力不足^[9],难以在微小故障智能诊断方面发挥更优越的性能。

1.2 改进的 CNNs-SVM 智能诊断算法

针对传统 CNN 网络的不足,本文将 CNN 强大的特征提取能力的优点与 SVM 在多分类功能的优越性相结合,提出了一种基于改进 CNNs-SVM 的新方法用于电机轴承的微小故障智能诊断,该方法在整个故障诊断过程中无需对原始数据做任何的手工特征提取和特征变换,只需将原始故障数据输入 CNNs-SVM 模型,诊断结果自动输出,“端到端”的算法结构具有更好的可操作性和通用性,其系统结构如图 2 所示。在图 2 中包含 3 个模块:最底层为故障数据采集模块,最上层为故障诊断的结果输出模块,中间层为 CNNs-SVM 算法模块。在 CNNs-SVM 算法结构中主要由输入层、特征提取层和分类判别输出层组成,其中特征提取层是 CNNs-SVM 算法的核心部分,包含多个依次循环堆叠的卷积层、激活层、池化层。

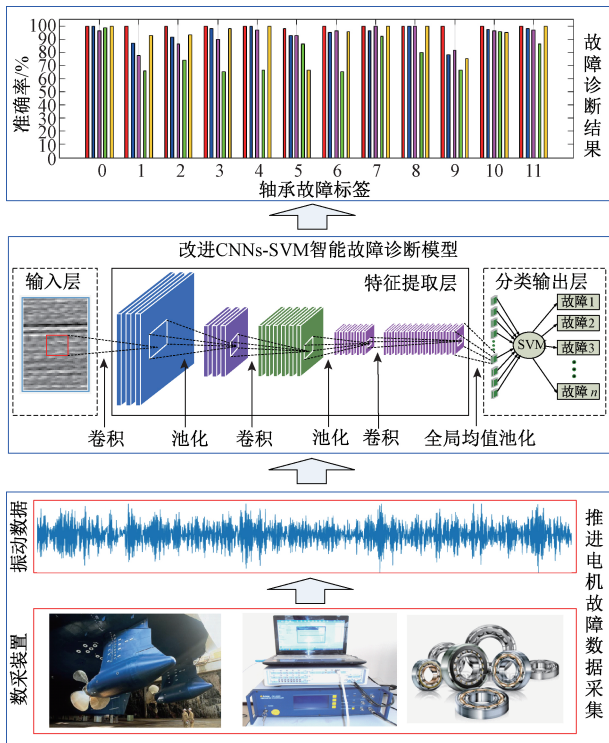


图 2 CNNs-SVM 自适应智能故障诊断

Fig.2 CNNs-SVM adaptive intelligent fault diagnosis

不同于传统的 CNN,改进的 CNNs-SVM 算法中设计了 1×1 的过渡卷积层和全局均值池化层^[23]的组合代替传统 CNN 中 2~3 层的全连接层部分;然后在全局均值池化层之后再并行连接 Softmax 层和 SVM 分类器,用于最终的故障分类结果输出。CNNs-SVM 算法的基本逻辑流程如图 3 所示。在图 3 中包含训练过程和测试过程两个阶段。在模型训练过程中,首先将 Softmax 函数与 CNNs

组合,通过 Softmax 函数进行误差反向传播最小化损失函数 $J(w)$,完成 CNNs 模型参数的训练;其次用训练好的 CNNs 模型对原始输入数据进行特征提取;然后再将提取到的低维特征向量和对应标签输入给 SVM 完成训练。在模型的测试阶段,同样采用已训练好的 CNNs 模型对新故障数据进行特征提取,将新故障输入数据转化为稀疏的低维特征向量,最后把提取到的低维特征数据输入给已训练好的 SVM 完成最终的故障分类,直接输出故障诊断的结果。

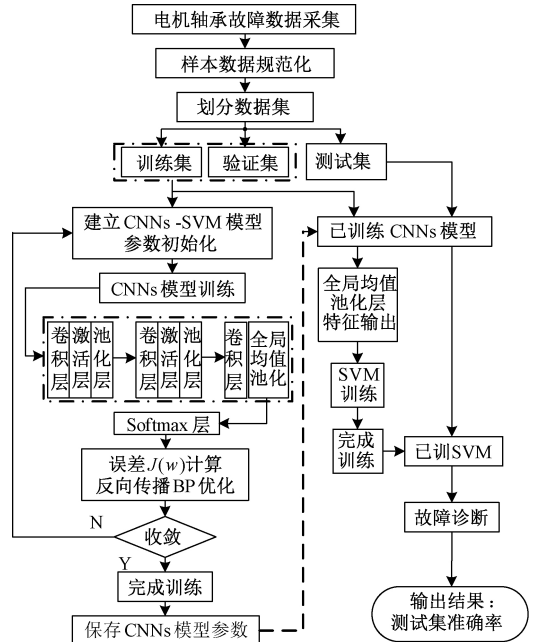


图 3 CNNs-SVM 智能故障诊断算法流程

Fig.3 The flow chart of the CNNs-SVM intelligent fault diagnosis algorithm

本文提出的 CNNs-SVM 算法基本组成单元如下。

1) 卷积层(convolution layer, CL)

卷积是一种对两个实变函数的数学运算,应用于机器学习可有效提取图片的特征信息^[20]。为了提取输入特征图的不同特征,需要使用多个不同的卷积核进行卷积操作。卷积操作的数学形式为:

$$y_i = f(x_{i+1}) = f(W_i \otimes x_i + b_i) \quad (2)$$

式中: x_i 为当前层的输入特征图; x_{i+1} 表示卷积计算后得到的特征图; $f(\cdot)$ 表示激活函数; y_i 表示经过激活层后的输出特征图; $\otimes$ 表示卷积运算符; W_i 表示卷积核权重; b_i 为偏置。激活函数有数十种^[5],在 CNN 中常用的激活函数有 Sigmoid、Tanh 和 Relu 激活函数;选用不同的激活函数会对模型的性能产生较大的差异性影响。

2) 池化层(pooling layer, PL)

PL 通常设置于 CL 之后,CL 为提升表达能力主要依靠增加提取特征的滤波器把原样本中尽可能多的潜在特

征提取出来^[20],池化操作就是在保证原有主要特征不丢失的基础上降维减参^[4]。池化操作主要有最大池化和均值池化^[19]。在 CNN 中最大池化最为常用,其数学表达式为:

$$\text{max-pooling}(f[i-1], f[i], f[i+1]) = \max(f[i-1], f[i], f[i+1]) \quad (3)$$

式中: max-pooling 表示最大池化; $f[i]$ 为第 i 个像素值; $f[i-1]$ 为第 $i-1$ 个像素值; $f[i+1]$ 为第 $i+1$ 个像素值; $\max(\cdot)$ 表示取最大值。

3) 全局均值池化层(global average pooling, GAP)

GAP^[23]是用来解决全连接网络参数量过多而提出的新技术,与池化操作相类似,GAP 是对最后一层卷积操作输出的特征图取全局平均值,其数学表达式为:

$$S_{\text{avg-pooling}}^l = \frac{1}{c} \sum_{i=1}^c X_{1:h,1:w,i}^l \quad (4)$$

式中: $S_{\text{avg-pooling}}^l$ 表示第 l 层特征图经全局均值池化得到的结果; $X_{1:h,1:w,i}^l$ 表示均值池化核所对应的输出特征图范围为横向方向从第 1 ~ h 行的像素点、纵向从第 1 ~ w 列的像素点。

4) Softmax 层

Softmax 函数是 Logistic 分类器的推广,主要用于多分类问题^[18]。假设训练输入样本为 x ,对应的标签为 $\bar{y}$,则将样本 x 判定为类别 j 的概率为 $P(y=j|x)$ 。所以,对于一个 n 分类的 softmax 分类器,输出是一个 n 维的向量(向量的各元素之和为 1,个体取值范围为 0~1 之间),softmax 函数^[18]的数学表达式为:

$$Y_i' = \begin{bmatrix} P(y_i = 1 | x_i) \\ P(y_i = 2 | x_i) \\ \vdots \\ P(y_i = n | x_i) \end{bmatrix}^T = \text{soft max}(Y_i) = \frac{1}{\sum_{l=1}^n e^{x_i^T \cdot w_l}} \cdot \begin{bmatrix} e^{x_i^T \cdot w_1} \\ e^{x_i^T \cdot w_2} \\ \vdots \\ e^{x_i^T \cdot w_n} \end{bmatrix}^T \quad (5)$$

式中: $w_1, w_2, \dots, w_n$ 为模型的训练参数; $P(y_i = 1 | x_i)$ 表示第 i 个样本 x_i 属于类别 1 的概率值; $1 / \sum_{l=1}^n e^{x_i^T \cdot w_l}$ 为归一化函数。

将归一化后的预测结果与对应样本标签相对比,即可计算误差损失值。在 CNN 中,常用的损失函数有均方误差损失函数^[20]和交叉熵损失函数^[10],由于本文的输入标签是分类标志,因此选用更适于位矢量的交叉熵损失函数^[4],如式(6)所示:

$$J(w) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^n I\{\bar{y}_i = j\} \log \frac{e^{x_i^T \cdot w_j}}{\sum_{l=1}^n e^{x_i^T \cdot w_l}} \right] \quad (6)$$

式中: i 表示第 i 个训练样本; j 表示属于第 j 个类别(共有 n 个类别); $I\{\cdot\}$ 为逻辑指示函数,当大括号内的值为真时, $I = 1$, 否则 $I = 0$; $\bar{y}_{(i)}$ 表示第 i 个样本的真实标签; $e^{x_i^T \cdot w_j} / \sum_{l=1}^n e^{x_i^T \cdot w_l}$ 表示第 i 个样本属于 j 类的 softmax 归一化概率值, $J(w)$ 为交叉熵损失函数;训练的过程即是不断的调整式(6)中的参数,以最小化损失函数 $J(w)$ 。

5) SVM 分类判别层

SVM 在解决非线性多分类任务时采用非线性 SVM 完成^[13]。通过使用核技巧和软间隔最大化,将线性不可分问题从原输入空间向高维特征空间映射,并在高维特征空间里构建最大间隔超平面,再采用线性分类算法求解非线性分类问题^[24]。滚动轴承的故障数据通常是线性不可分的训练数据,假设训练数据集为 $Z = \{(x_1, y_1), \dots, (x_n, y_n)\}$, $x_i \in X = R^n$, $y_i \in \psi = \{-1, +1\}$, $i = 1, \dots, n$; x_i 为第 i 个特征向量, y_i 为 x_i 的类别标记。因此,可对训练样本 (x_i, y_i) 引进松弛变量 ξ_i ,从而使线性不可分的线性 SVM 的学习问题描述为软间隔最大化问题^[25]:

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & y_i(w \cdot x_i - b) \geq 1 - \xi_i \quad \forall (x_i, y_i) \in Z \\ & \xi_i \geq 0, i = 1, 2, \dots, N \end{aligned} \quad (7)$$

式中: w 和 b 为优化参数。通过调整 w 和 b ,使得目标函数最小化,从而得到最大间隔分离超平面^[13]。通过求解式(7)的对偶问题并选取适当的正定核函数 $K(x_i \cdot x_j)$ 和惩罚参数 C ($C > 0$),即可得到求解非线性问题的分离超平面数学表达式:

$$\sum_{i=1}^n \alpha_i^* y_i K(x \cdot x_i) + b^* = 0 \quad (8)$$

式中: α_i^* 和 b^* 为最优解。本文的 CNNs-SVM 方法中,采用非线性 SVM 对 CNNs 提取的特征数据进行训练,核函数 $K(x \cdot z)$ 采用了高斯径向基核函数(radial basis function, RBF)^[24],使用 RBF 核函数后的分类决策函数可表达为:

$$f(x) = \text{sign} \left(\sum_{i=1}^n \alpha_i^* y_i \exp \left(-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right) + b^* \right) \quad (9)$$

2 电机轴承故障诊断实验验证

为评估提出的改进 CNNs-SVM 方法对故障诊断的有效性和准确性,本文以滚动轴承为对象,运用滚动轴承在不同工况、不同运行状态下的振动信号数据建立故障数据集,然后用于 CNNs-SVM 模型的训练和测试,完成实验验证。

2.1 故障实验数据集

本实验数据来自于美国凯斯西储大学(CWRU)的滚

动轴承实验台,如图 4 所示^[19]。被测试的故障轴承型号为 6205-2RS JEM,用于支撑电动机主轴。本文实验中采用电火花方法分别在多个滚动轴承的外圈、内圈和滚动体表面上依次精加工出直径为 0.007、0.014、0.021 和 0.028 inches (1 inches≈25.4 mm) 的单点凹坑,用于模拟轴承在实际工作中最易出现的点蚀失效形式,尺寸的不同表示点蚀故障从微小故障逐渐增大。本实验的采样频率为 12 kHz,采样时间为 10 s,分别采集了 0 马力 (1 797 r/min)、1 马力 (1 772 r/min)、2 马力 (1 750 r/min) 和 3 马力 (1 730 r/min) 4 种负载状态下的轴承振动加速度数据^[26]。

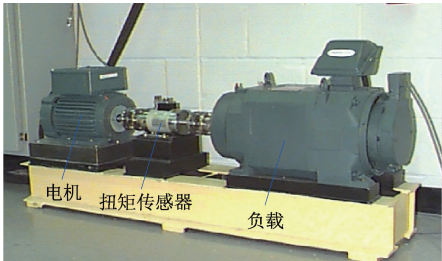


图 4 滚动轴承故障数据采集实验台

Fig.4 The rolling bearing fault data acquisition experiment bench

根据以上工作,制作用于 CNNs-SVM 模型训练的样本数据集。在本实验中,4 种负载下的转速范围为 1 730~1 797 r/min,12 kHz 的采样频率下每秒采集 12 000 个点。转轴每转一圈传感器采集的点数范围为 400~416 个点 (12 000×60/1 797≈400),为保证故障数据的可信度,每类故障的每个样本长度设置为 500 个采样点。本文选取了滚动轴承在 0 马力负载下的 12 种轴承工作状态数据作为训练数据,其中包含 11 种故障类型和 1 个正常状态,每种工作状态包含 200 个样本,每个样本包含 500 个振动数据点,轴承故障数据集描述如表 1 所示。

2.2 故障数据处理

为更好地训练改进的 CNNs-SVM 模型,对表 1 所示的滚动轴承数据集进行必要的数据处理。

1) 输入数据格式重构。表 1 所示的数据集为一维时间序列数据,为更好地实现输入特征图表达,便于有效的卷积和池化操作,本文将一维原始数据 [500,1] 重构为二维输入特征图形式 [25,20,1],其中 1 代表 1 通道。重构方法^[27]如图 5 所示,具体过程为:首先将每个长度为 500 的一维时间序列样本进行等长截取,截取的份数为 25 段,每个数据段为 20 个数据点,然后将 25 个数据段堆叠,从而可获得一个 25×20 的二维特征图。因此,每类故障状态均有 200 个格式为 [25, 20, 1] 的输入特征图样本数据。

表 1 滚动轴承故障数据集
Table 1 The rolling bearing fault dataset

标签	故障位置	故障尺寸/mm	故障程度	样本长度	样本数量
0	正常	无	无	500	200
1	滚球	孔径:0.18,坑深:0.28	微小	500	200
2	滚球	孔径:0.36,坑深:0.28	轻度	500	200
3	滚球	孔径:0.53,坑深:0.28	中度	500	200
4	滚球	孔径:0.71,坑深:0.28	显著	500	200
5	内圈	孔径:0.18,坑深:0.28	微小	500	200
6	内圈	孔径:0.36,坑深:0.28	轻度	500	200
7	内圈	孔径:0.53,坑深:0.28	中度	500	200
8	内圈	孔径:0.71,坑深:1.27	显著	500	200
9	外圈	孔径:0.18,坑深:0.28	微小	500	200
10	外圈	孔径:0.36,坑深:0.28	轻度	500	200
11	外圈	孔径:0.53,坑深:0.28	中度	500	200

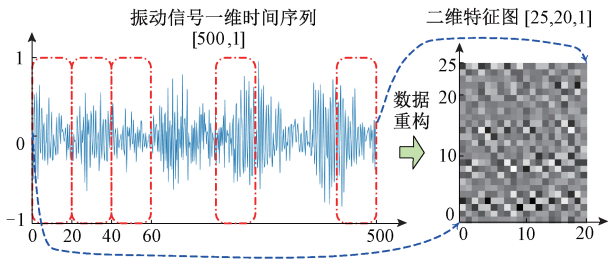


图 5 输入数据格式重构

Fig.5 The data format reconstruction of input data

2) 对输入特征图数据进行标准化。为了更好地训练 CNN 网络模型,对每个样本输入数据做数据标准化处理,标准化数学表达式^[20]为:

$$X = \{x_i\} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

(10)

3) 数据集划分。将标准化处理后的 200 个样本数据随机的划分为训练数据集和测试数据集,比例分别为 70% 和 30%,在训练集中随机取 20% 作为验证数据集^[4]。首先将训练集用于 CNNs-SVM 模型的训练,训练过程中根据验证集精度曲线调整模型超参数;然后将验证集用于训练后的 CNNs-SVM 模型准确性验证;最后将测试集用于模型最终的故障分类精度评估。

2.3 CNNs 模型结构的选择

CNNs-SVM 模型训练是故障诊断的关键环节,模型结构超参数的选择对故障诊断准确性的影响重大^[19]。本文为考察不同层数的 CNN 结构对准确率性能的影响,分别设计了 3 种 CNN 拓扑结构,每一种拓扑结构都包含

了普通训练和加入深度学习训练技巧^[25]两种情况,其超参数如表 2 所示。N1、N2、N3 分别为使用了 1 组、2 组、3 组 CNN 模型堆叠在一起;S1~S3 表示在 N1~N3 模型中添加了深度学习训练技巧后的模型,训练技巧包含有数据增强^[28]和 Dropout^[20]。经过大量反复对比计算,本文

优选 Relu 函数作为 CNN 的激活函数,采用 Tanh 函数作为全连接层的激活函数,采用了 Adam 自适应优化器,经试验 mini-batch 设置为 64,循环迭代轮数为 100 轮。实验结果如表 3 所示,6 种模型迭代 100 轮的准确率曲线如图 6 所示。

表 2 不同 CNN 拓扑结构模型的超参数
Table 2 The hyper-parameters of different CNN topology structure models

类型	序号	CNN 模型超参数
普通训练	N1	卷积(64,2)+池化(2,2)+全连接(128/12)
	N2	卷积(64,2)+池化(2,2)+卷积(32,2)+池化(2,2)+全连接(128/12)
	N3	卷积(64,2)+池化(2,2)+卷积(64,2)+池化(2,2)+卷积(32,2)+池化(2,2)+全连接(128/12)
增加技巧	S1	数据增强+卷积(64,2)+池化(2,2)+全连接(128/12)+Dropout(0.5)
	S2	数据增强+卷积(64,2)+池化(2,2)+卷积(32,2)+池化(2,2)+全连接(128/12)+Dropout(0.5)
	S3	数据增强+卷积(64,2)+池化(2,2)+卷积(64,2)+池化(2,2)+卷积(32,2)+池化(2,2)+全连接(128/12)+Dropout(0.5)

表 3 不同 CNN 模型的故障识别准确率
Table 3 The fault identification accuracies of different CNN models

模型结构	测试准确率/%		
	一层 CNN	二层 CNN	三层 CNN
普通训练	87.64	97.28	94.76
技巧训练	93.72	98.75	96.11

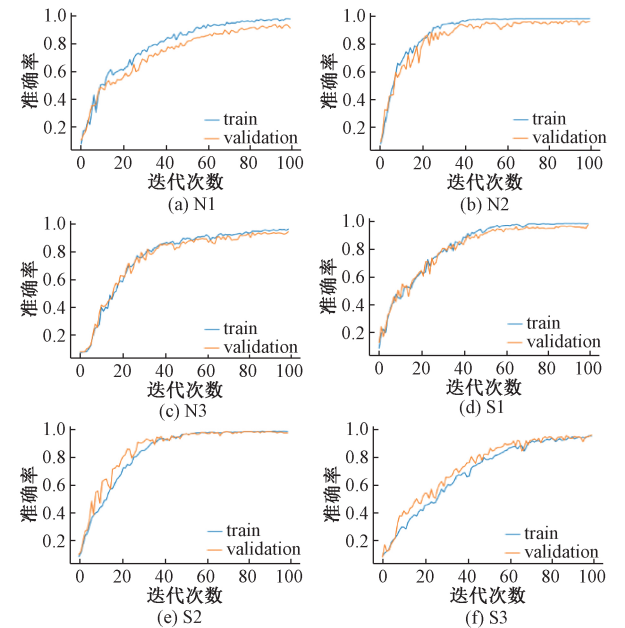


图 6 6 种 CNN 模型训练性能比较

Fig.6 The training performance comparison of 6 CNN models

从表 3 和图 6 可以明显看出,使用训练技巧后的两层 CNN 模型测试准确率最好,已达到 98.75%,并且有效

改善了过拟合问题。从图 6(b)可以看出,未使用训练技巧的训练准确率虽然接近 99%,但是测试准确率仅为 97.28%,已出现过拟合;普通训练的一层 CNN 网络由于结构较简单,数据集特征提取力不足,准确率仅为 87.64%;使用了三层的 CNN 反而不如两层的 CNN 结构,原因在于本实验的轴承数据样本量较少,从而导致深层的网络结构容易出现梯度消失和过拟合问题。对比图 6(a)~(c)与(d)~(f)可以看出采用了数据增强和 dropout 训练技巧后可以有效改善过拟合问题,同时也在小范围内提升了模型的准确率。

3 结果分析与对比验证

3.1 CNNs-SVM 的诊断结果

根据以上研究的结果和结论,最终建立 CNNs-SVM 故障智能诊断模型,具体超参数如表 4 所示。该模型在表 2 的第 S2 组两层的 CNN 模型基础上进行改进,每一组 CNN 后设置 Dropout 层,在两层 CNN 之后设置了一个只改变输出维度不改变特征图形状的 1×1 的 CL,用于调整输出维度为 12,对应 12 种故障类型;然后 GAP 代替全连接层结构放置于 1×1 的 CL 之后进行高效降维减参。采用表 2 第 S2 组全连接层的 CNN 结构的总参数量为 92 140 个,而 CNNs-SVM 的总训练参数量仅为 9 096 个,如表 5 所示,参数量计算方法详见文献[20]。

在 CNNs-SVM 模型训练中同样采用 Adam 自适应学习率优化算法,采用数据增强和 Dropout 改善过拟合问题,采用 Relu 激活函数,mini-batch 为 64,循环迭代轮数为 100 轮,SVM 参数设置为惩罚系数 $C=10$,采用高斯径向基核函数(RBF),松弛变量 $\xi=0.1$ 。最终 CNNs-SVM

表 4 提出的故障诊断模型超参数

Table 4 The hyper-parameters of the proposed fault diagnosis model

CNNs-Softmax 训练环节网络模型			
网络层/名称	结构参数	输出形状	参数量
输入层	[batch , 25 , 20 , 1]	[batch , 25 , 20 , 1]	0
CL 1	[2,2,64,1], “Valid”, 步长=[1,1]	[batch , 24 , 19 , 64]	320
激活层 1	Relu 激活函数	[batch , 24 , 19 , 64]	0
PL 1	[1,2,2,1], “Same”, 步长=[2,2]	[batch , 12 , 9 , 64]	0
Dropout 层 1	Dropout (0. 2)	[batch , 12 , 9 , 64]	0
CL 2	[2,2,64,32], “Valid”, 步长=[1,1]	[batch , 12 , 9 , 32]	8 224
激活层 2	Relu 激活函数	[batch , 12 , 9 , 32]	0
PL 2	[1,2,2,1], “Same”, 步长=[2,2]	[batch , 6 , 4 , 32]	0
Dropout 层 2	Dropout (0. 1)	[batch , 6 , 4 , 32]	0
CL 3	[1,1,32,12], “Same”, 步长=[1,1]	[batch , 6 , 4 , 12]	396
GAP	[1,5,4,1], “Same”, 步长=[5,4]	[batch , 12]	0
CNN 特征输出层	CNN 模型参数已保存	[batch , 12]	0
中间输出层	Softmax 激活函数	[batch , 12]	0
CNNs-SVM 测试环节网络模型			
输入层	[batch , 25 , 20 , 1]	[batch , 25 , 20 , 1]	0
CNN 特征输出层	将原始数据输入已训练好的 CNNs 直接提取特征	[batch , 12]	0
SVM 分类判别层	非线性 SVM 分类器	[batch , 12]	156
最终输出层	故障分类结果	[batch , 12]	0

表 5 模型训练参数对比表

Table 5 The model training parameter comparison table

名称	CNNs-全连接	CNNs-SVM
CL 1	320	320
CL 2	8 224	8 244
CL 3	无	396
全连接层 1	82 048	无
全连接层 2	1 548	无
SVM 分类器	无	156
总参数量	92 140	9 096

在测试集上的最高准确率为 99. 86%, 其诊断结果如表 6 所示。

表 6 不同的 CNN 模型的故障诊断结果

Table 6 The fault diagnosis results of different CNN models

名称	测试准确率/%	训练时间/s	测试时间/s
改进的 CNNs+SVM	99. 86	230. 116	0. 187
改进的 CNNs+Softmax	99. 04	229. 53	0. 198
传统全连接 CNNs+Softmax	98. 75	248. 96	0. 279

本文设计的 CNNs-SVM 算法包含图 3 所示的两个部分,从而表 6 中包含两个结果,其一是改进的 CNNs 与 Softmax 组合进行反向传播优化训练 CNNs 模型时得到的准确率,另一个是将训练好的 CNNs 模型作为特征提取器对新故障数据进行特征提取后再输入 SVM 进行故障分类得到的准确率。对比表 6 可知,在准确率方面:CNNs+Softmax 部分的测试集准确率为 99. 04%,而改进的 CNNs-SVM 的准确率已提升至 99. 86%;在时间方面:CNNs 模型的训练时间为 229. 53 s, SVM 的训练时间为 0. 586 s, 即 CNNs-SVM 总的训练时间为 230. 116 s,可见本文方法虽然在 CNNs 的基础上多加了 SVM 部分,但是训练时间并未过多增加。另外,对比表 6 可见,改进的 CNN 性能明显优于传统的全连接 CNN,尤其在测试时间方面有极大的提升,这对提出的方法应用于微小故障的快速诊断与在线监测具有重要意义。

3. 2 CNNs-SVM 的结果评估

为了更有效地评估 CNNs-SVM 算法的准确性,本文引用了精确率^[5]和召回率^[28]对 CNNs-SVM 算法进行评估。限于篇幅,本文仅以表 6 中 CNNs-SVM 的实验结果计算得到的精确率和召回率,如表 7 所示。

表 7 改进的 CNNs-SVM 模型诊断结果评估

Table 7 The diagnosis result evaluation of the improved CNNs-SVM model

故障序号	精确率	召回率	F1 调和均值	样本
0	1.000 0	1.000 0	1.000 0	60
1	1.000 0	0.983 3	0.991 6	60
2	1.000 0	1.000 0	1.000 0	60
3	1.000 0	1.000 0	1.000 0	60
4	1.000 0	1.000 0	1.000 0	60
5	1.000 0	1.000 0	1.000 0	60
6	1.000 0	1.000 0	1.000 0	60
7	1.000 0	1.000 0	1.000 0	60
8	1.000 0	1.000 0	1.000 0	60
9	1.000 0	1.000 0	1.000 0	60
10	0.983 6	1.000 0	0.991 7	60
11	1.000 0	1.000 0	1.000 0	60
平均值/总数	0.998 6	0.998 6	0.998 6	720

为进一步考察 CNNs-SVM 算法对故障误判的细节情况,本文引入了多分类混淆矩阵^[24]对诊断结果进行量化分析。混淆矩阵比较全面的反映了滚动轴承在不同故障等级下的诊断正确率和误判数量,以及真实故障类型被误判为何种类型等信息,滚动轴承对应表 7 的混淆矩阵如图 7 所示。

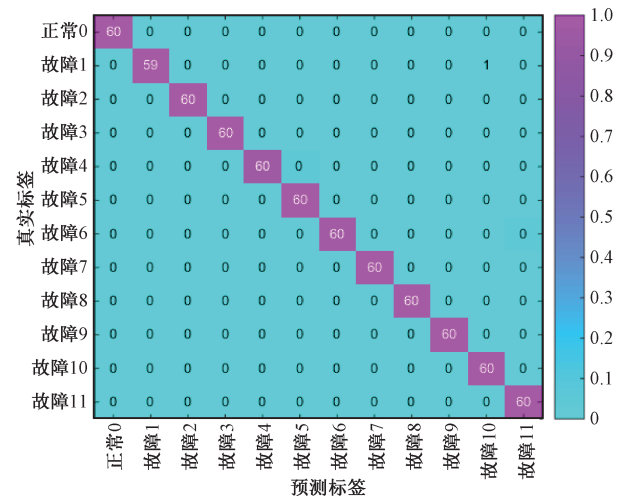


图 7 故障诊断多分类混淆矩阵

Fig.7 Multi-class confusion matrix of fault diagnosis results

在图 7 中, x 轴横坐标代表故障的预测类别, y 轴纵坐标代表故障的真实标签;每一类故障的测试集样本数为 60 个,共 12 类工作状态;主对角线上的数字代表了 CNNs-SVM 算法对每一类故障状态正确诊断的样本个数。从图 7 中明显可见:在测试集上的 720 个样本中只有 1 个样本被预测错误,被错判样本的真实标签为故障

类型 1(滚珠的微小故障),而预测类别为故障 10(外圈的轻度故障);除此之外其他类型故障的诊断准确率均为 100%。由此可见,CNNs-SVM 算法对微小故障定位识别的准确率高达为 100%,且区分故障与正常状态的识别率为 100%,综合故障识别率高达 99.86%。经实验验证,CNNs-SVM 算法对滚动轴承的微小故障具有优越的识别能力和较高的诊断准确率。

3.3 与其他算法对比分析

为验证本文提出的方法相比当前主流智能诊断方法的有效性,本文将 CNNs-SVM 诊断结果与目前主流的 SVM、BPNN、KNN 及深层的 BP 神经网络(DNN)进行对比验证。在传统的智能故障诊断方法中通常先对原始数据集进行手工特征提取,再将提取的特征输入智能诊断算法完成分类预测^[4,6]。Xia 等^[4]采用了 14 种方法对原始数据进行特征提取前处理,包含 10 种时域特征提取方法和 4 种频域特征提取方法。本文按照同样的方法对表 1 所示的 12 类轴承状态数据提取其波峰、峭度、绝对平均值、均方根和方差值等特征指标^[4],然后再将提取的特征分别用于 SVM、BPNN、KNN 和 DNN 算法进行故障诊断,实验的结果如表 8 和 9 所示。

表 8 5 种诊断算法耗时对比

	methods (s)				
算法	CNNs-SVM	SVM	KNN	BPNN	DNN
训练时间/s	230.116	2.264	0.049	136.32	327.09
测试时间/s	0.187	0.829	1.322	0.082	0.141

表 9 5 种算法诊断准确率数据表

	The precision rate of five diagnosis methods (%)				
故障类别	CNNs-SVM	SVM	KNN	BPNN	DNN
0	100	100	96.25	98.89	100.00
1	100	96.72	78.00	66.15	92.86
2	100	92.01	86.31	74.07	93.55
3	100	98.31	90.15	65.56	98.36
4	100	100.00	97.00	66.67	100.00
5	100	93.25	92.85	86.49	66.59
6	100	95.24	96.59	65.17	96.08
7	100	87.32	100.00	92.31	100.00
8	100	100.00	100.00	80.01	100.00
9	100	78.59	81.69	66.41	75.56
10	98.36	97.78	96.72	95.83	95.24
11	100	98.18	97.00	86.30	100.00
平均值	99.86	94.78	92.71	78.66	93.19

对比表 9 可以明显看出,本文所提出的 CNNs-SVM 算法的准确率明显优于其他算法。经过手工特征提取后的 SVM、KNN、BPNN 精确率分别为 94.78%、92.71%、78.66%。具有 4 个隐含层的 DNN 对原始数据直接训练的精确率为 93.19%;只有一个隐含层的 BP 网络训练准确率很低,而且过拟合严重;DNN 当层数超过 5 层时,受梯度消失影响准确率难以提升且过拟合突出。可以看出,5 种算法对正常状态识别的准确率均较高,然而在微小故障识别方面,CNNs-SVM 算法明显优于其他算法。通过对比以上 5 种算法结果可以看出,提出的 CNNs-SVM 算法具有更加优越的效果。

3.4 不同负载的迁移实验

为了评估 CNNs-SVM 算法应对负载变化时故障诊断的迁移通用性能力,本文运用 3.1 节所述的相同方法再次对轴承实验台在 1 和 2 马力两种负载下采集的 12 类故障数据进行训练和测试,迁移实验所使用的 CNNs-SVM 模型超参数与 0 马力实验完全相同,实验结果如表 10 所示,并使用混淆矩阵对两次迁移实验结果进行量化,如图 8 所示。

表 10 迁移实验故障诊断结果

Table 10 Fault diagnosis results of migration experiment

负载	1 马力		2 马力	
	精确率	召回率	精确率	召回率
0	1.000 0	1.000 0	1.000 0	1.000 0
1	1.000 0	1.000 0	1.000 0	0.983 3
2	1.000 0	1.000 0	1.000 0	0.966 7
3	1.000 0	1.000 0	1.000 0	0.983 3
4	1.000 0	0.966 7	0.937 5	1.000 0
5	0.967 7	1.000 0	0.967 2	0.983 3
6	1.000 0	0.983 3	1.000 0	1.000 0
7	1.000 0	1.000 0	1.000 0	1.000 0
8	1.000 0	1.000 0	1.000 0	1.000 0
9	1.000 0	1.000 0	1.000 0	1.000 0
10	1.000 0	1.000 0	0.983 1	0.966 7
11	0.983 6	1.000 0	1.000 0	1.000 0
平均值	0.995 9	0.995 8	0.990 6	0.990 3

对比表 10 可知,改进的 CNNs-SVM 算法对滚动轴承在 1 和 2 马力两种负载下的迁移实验中仍获得了较高的故障识别精确率,分别为 99.59%和 99.06%,充分说明本文所提出的 CNNs-SVM 方法具有较好的迁移通用性。从图 8 所示的两次负载实验的混淆矩阵可以看出,两者对正常状态识别的准确率均为 100%;在 1 马力实验中的 720 个样本中,只有 3 个样本发生错判,准确率接近

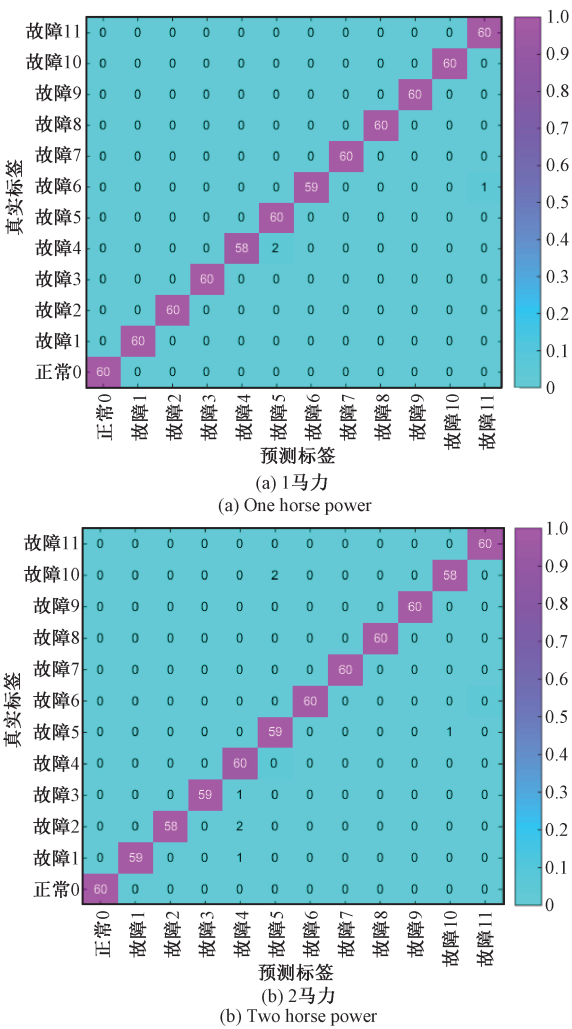


图 8 故障诊断迁移实验的多分类混淆矩阵

Fig.8 Multi-class confusion matrix of fault diagnosis migration test

99.6%;在 2 马力实验中有 7 个样本被误判,准确率也高达 99%。以上误判均为不同故障类别之间的差错,然而在正常状态与故障状态两种工作状态之间的识别准确率均为 100%。以上实验表明,本文提出的 CNNs-SVM 算法具有良好的迁移通用性能,在不同负载工况下同样取得了较高的故障识别率并可对故障发生位置准确定位,这对滚动轴承在实际工作中故障的快速定位诊断与定点维护以及故障多发部件优化改进有着重要的指导意义。

4 结 论

针对传统 CNN 与现在智能诊断算法的不足,提出了基于改进 CNNs-SVM 的深度学习新算法用于电机轴承微小故障的智能化诊断。通过对滚动轴承实验数据的故障诊断,验证了该方法的准确性与优越性,结论如

下:改进的 CNNs-SVM 算法具有更高的故障诊断准确率和更快的诊断速度,在滚动轴承故障实验中获得了高达 99.86% 识别准确率,明显优于传统的基于手工特征提取的 SVM、KNN、BPNN 和 DNN 的准确率,且在微小故障诊断方面具有更优越的性能;通过混淆矩阵可直接量化故障误判的位置和数量。CNNs-SVM 具有更短的测试时间,更有利于实现在线监测和快速智能诊断的应用。

采用全局均值池化技术和 1×1 的 CL 组合代替传统 CNN 全连接层结构,可有效减少传统 CNN 的训练参数量;使用数据增强、Dropout 和批量归一化等深度学习训练技巧,可以有效防止模型过拟合,并提高了算法的鲁棒性和迁移能力。通过将 CNNs-SVM 模型应用于滚动轴承在 1 和 2 马力两种负载数据集,分别得到了 99.59% 和 99.07% 的故障识别准确率。CNNs-SVM 算法无需对原始故障数据做任何手工特征提取,直接以原始数据作为模型输入,故障诊断结果自动输出,端到端的算法结构具有更好的可操作性和通用性。

参考文献

[1] 马伟明. 舰船综合电力系统中的机电能量转换技术[J]. 电气工程学报, 2015, 10(4): 3-10.
MA W M. Electromechanical power conversion technologies in vessel integrated power system [J]. Journal of Electrical Engineering, 2015, 10(4): 3-10.

[2] KARMAKER H, SARANDRIA D, MAN T H, et al. High-power dense electric propulsion motor [J]. IEEE Transactions on Industry Applications, 2015, 51(2): 1341-1347.

[3] LIU R N, YANG B Y, ZIO E, et al. Artificial intelligence for fault diagnosis of rotating machinery: A review [J]. Mechanical Systems and Signal Processing, 2018, 108:33-47.

[4] XIA M, LI T, XU L, et al. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks [J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(1): 101-110.

[5] SHAO H, JIANG H, LIN Y, et al. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders [J]. Mechanical Systems and Signal Processing, 2018, 102: 278-297.

[6] 温江涛, 闫常弘, 孙洁娣, 等. 基于压缩采集与深度学习的轴承故障诊断方法[J]. 仪器仪表学报, 2018, 39(1): 171-179.
WEN J T, YAN CH H, SUN J D, et al. Bearing fault diagnosis method based on compressed acquisition and deep learning [J]. Chinese Journal of Scientific Instrument, 2018, 39(1): 171-179.

[7] HSIAO T, WENG M C. A hierarchical multiple-model

approach for detection and isolation of robotic actuator faults [J]. Robotics and Autonomous Systems, 2012, 60(2): 154-166.

[8] AMAR M, GONDAL I, WILSON C C. Vibration spectrum imaging: A novel bearing fault classification approach [J]. IEEE Transactions on Industrial Electronics, 2015, 62(1): 494-502.

[9] 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 CH L, LYU F Y, BAO ZH J, et al. A review of data driven-based incipient fault diagnosis [J]. Acta Automatica Sinica, 2016, 42(9): 1285-1299.

[10] 李恒, 张氢, 秦仙蓉, 等. 基于短时傅里叶变换和卷积神经网络的轴承故障诊断方法[J]. 振动与冲击, 2018, 37(19): 124-131.
LI H, ZHANG Q, QIN X R, et al. Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network [J]. Journal of Vibration and Shock, 2018, 37(19): 124-131.

[11] JEGADEESHWARAN R, SUGUMARAN V. Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines [J]. Mechanical Systems and Signal Processing, 2015, 52-53(1): 436-446.

[12] YANG Y, YU DEJIE, CHENG J. A roller bearing fault diagnosis method based on EMD energy entropy and ANN [J]. Journal of Sound and Vibration, 2006, 294(1): 269-277.

[13] 姚德臣, 杨建伟, 程晓卿, 等. 基于多尺度本征模态排列熵和 SA-SVM 的轴承故障诊断研究[J]. 机械工程学报, 2018, 54(9): 168-176.
YAO D CH, YANG J W, CHENG X Q, et al. Railway rolling bearing fault diagnosis based on multi-scale IMF permutation entropy and SA-SVM classifier [J]. Journal of Mechanical Engineering, 2018, 54(9): 168-176.

[14] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-507.

[15] 李巍华, 单外平, 曾雪琼. 基于深度信念网络的轴承故障分类识别[J]. 振动工程学报, 2016, 29(2): 340-347.
LI W H, SHAN W P, ZENG X Q. Bearing fault identification based on deep belief network [J]. Journal of Vibration Engineering, 2016, 29(2): 340-347.

[16] JIA F, LEI Y, LIN J, et al. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive

- data [J]. *Mechanical Systems and Signal Processing*, 2016, 72-73: 303-315.
- [17] TAMILSELVAN P, WANG P. Failure diagnosis using deep belief learning based health state classification [J]. *Reliability Engineering and System Safety*, 2013, 115(7): 124-135.
- [18] 王丽华, 谢阳阳, 周子贤, 等. 基于卷积神经网络的异步电机故障诊断[J]. *振动、测试与诊断*, 2017(6): 1208-1215.
- WANG L H, XIE Y Y, ZHOU Z X, et al. Asynchronous motor fault diagnosis based on convolutional neural network [J]. *Journal of Vibration, Measurement and Diagnosis*, 2017(6): 1208-1215.
- [19] 曲建岭, 余路, 袁涛, 等. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法[J]. *仪器仪表学报*, 2018, 39(7): 134-143.
- QU J L, YU L, YUAN T, et al. Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network [J]. *Chinese Journal of Scientific Instrument*, 2018, 39(7): 134-143.
- [20] GOODFELLOW I, BENGIO Y, COURVILLE A. *Deep learning* [M]. The MIT Press, 2016.
- [21] LECUN Y, BENGIO Y, HINTON G. *Deep learning* [J]. *Nature*, 2015, 521: 436-444.
- [22] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]. *International Conference on Learning Representations* 2015, 2015, 1-14.
- [23] LIN M, CHEN Q, YAN S C. Network in network [C]. *International Conference on Learning Representations* 2014, 1-10.
- [24] FLACH P. *Machine Learning* [M]. Cambridge University Press, 2012.
- [25] HAN X, WANG J, WU Z, et al. Learning solutions to two dimensional electromagnetic equations using LS-SVM [J]. *Neurocomputing*, 2018, 317(23): 15-27.
- [26] 宫文峰, 黄美发, 张美玲, 等. 基于灵敏度及层次分析法的键合头多目标结构优化[J]. *振动与冲击*, 2015, 34(16): 128-134.

GONG W F, HUANG M F, ZHANG M L, et al. Multi-objective optimization of bonding head based on sensitivity and analytic hierarchy process [J]. *Journal of Vibration and shock*, 2015, 34(16): 128-134.

- [27] HOANG D T, KANG H J. Rolling element bearing fault diagnosis using convolutional neural network and vibration image [J]. *Cognitive Systems Research*, 2019, 53: 42-50.
- [28] MONTERRAT D M, LIN Q, ALLEBACH J, et al. Training object detection and recognition CNN models using data augmentation [J]. *Electronic Imaging*, 2017, 2017(10): 27-36.

作者简介



宫文峰, 2009 年于山东科技大学获得学士学位, 2014 年于桂林电子科技大学获得硕士学位, 现为武汉理工大学与新加坡南洋理工大学联合培养博士研究生, 主要研究方向为智能故障诊断与健康状态监测、深度学习与机器学习。

E-mail: wfgongcn@163.com

Gong Wenfeng received his B. Sc. degree in 2009 from Shandong University of Science and Technology and M. Sc. degree in 2014 from Guilin University of Electronic and Technology. Now, he is a Ph. D. candidate in Wuhan University of Technology and Nanyang Technological University, Singapore. His main research interest includes intelligent fault diagnosis and health condition monitoring, deep learning and machine learning.



陈辉(通信作者), 分别在 1984 年、1987 年和 1991 年于武汉理工大学获得学士学位、硕士学位和博士学位, 现为武汉理工大学博士生导师, 国家二级教授, 主要研究方向为船舶电力推进及智能船舶技术、电力系统建模与新能源技术等。

E-mail: hchen@whut.edu.cn

Chen Hui (Corresponding author) received his B. Sc., M. Sc. and Ph. D. degrees all from Wuhan University of Technology in 1984, 1987 and 1991, respectively. Now, he is a professor and doctoral supervisor in Wuhan University of Technology. His main research interest includes marine electric propulsion and intelligent ship technology, power system modeling and new energy source technology.