

西安交通大学学报
Journal of Xi'an Jiaotong University
ISSN 0253-987X, CN 61-1069/T

《西安交通大学学报》网络首发论文

题目：改进 CNN 的滚动轴承变工况故障诊断
作者：赵小强，张亚洲
收稿日期：2021-05-07
网络首发日期：2021-08-10
引用格式：赵小强，张亚洲. 改进 CNN 的滚动轴承变工况故障诊断. 西安交通大学学报.
<https://kns.cnki.net/kcms/detail/61.1069.T.20210810.1435.007.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

改进 CNN 的滚动轴承变工况故障诊断

赵小强^{1,2,3}, 张亚洲¹

(1.兰州理工大学 电气工程与信息工程学院, 兰州 730050; 2.甘肃省工业过程先进控制重点实验室, 兰州 730050;
3.兰州理工大学 国家级电气与控制工程实验室教学中心, 兰州 730050)

摘要: 针对滚动轴承在强噪声环境和变工况下故障诊断效果不佳、泛化能力差的问题, 提出一种改进 CNN 的滚动轴承变工况故障诊断方法。设计了多尺度特征提取模块, 采用不同尺度的卷积层对输入数据进行特征提取, 实现最大化提取故障数据中的特征信息, 同时引入通道注意力机制, 提取出该模块中更重要、更关键的信息; 设计了带跳跃连接线的卷积模块, 防止提取到的丰富特征在卷积层前向传递时丢失; 以 softmax 交叉熵作为损失函数, 利用 Adam 优化算法实现滚动轴承故障诊断。将所提方法分别在凯斯西储大学轴承数据集和变速箱数据集下进行实验验证, 结果表明: 在凯斯西储大学轴承数据集上的变噪声实验中, 所提方法诊断平均准确率为 96.49%, 在变工况中诊断准确率在 90% 以上, 均高于比较方法; 在变速箱轴承数据集上, 所提方法诊断准确率为 99.54%, 具有较好的抗噪性和更好的泛化能力。

关键词: 故障诊断; 滚动轴承; 变工况; 卷积神经网络; 注意力机制
中图分类号: TH133.3; TP206.3 **文献标识码:** A

Improved CNN Fault Diagnosis Method of Rolling Bearings under Variable Working Conditions

ZHAO Xiaoqiang^{1,2,3}, ZHANG Yazhou¹

(1.College of Electrical Engineering and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;

2.Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou 730050, China;

3.National Experimental Teaching Center of Electrical and Control Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Aiming at the problem of poor fault diagnosis and poor generalization ability of rolling bearing in a strong noise environment and variable working conditions, an improved CNN-based fault diagnosis method of rolling bearing variable working conditions is proposed. A multi-scale feature extraction module is designed, and convolutional layers of different scales are used to extract features from the input data to maximize the extraction of feature information in the fault data. At the same time, the channel attention mechanism is introduced to extract the more important and critical components of this module. Designed a convolution module with skip connection lines to prevent the extracted rich features from being lost when the convolutional layer is forwarded. Using softmax cross entropy as the loss function, the Adam optimization algorithm is used to realize the fault diagnosis of rolling bearing. The proposed method was verified by experiments on the bearing dataset and gearbox dataset of Case Western Reserve University. The results show that: in the variable noise experiment on the bearing dataset of Case Western Reserve University, the proposed method has an average diagnostic accuracy rate is 96.49%, and the diagnostic accuracy rate is over 90% in variable working conditions, which is higher than the comparison method. On the gearbox bearing data set, the diagnostic accuracy rate of the proposed method is 99.54%, which has better noise resistance and generalization ability.

收稿日期: 2021-05-07。 作者简介: 赵小强(1969-), 男, 教授, 博士生导师。 基金项目: 国家自然科学基金资助项目(61763029); 甘肃省高等学校产业支撑引导项目(2019C-05); 甘肃省工业过程先进控制重点实验室开放基金项目(2019KFJJ01)。

Keywords: fault diagnosis; rolling bearing; operation mode; convolutional neural network; attention mechanism

随着“智能制造”战略的不断推进,现代化机械设备开始朝着自动化、高效化和智能化趋势发展。旋转机械是机械设备中最重要的组成部分,被广泛应用于国民经济的各个行业中。滚动轴承作为旋转机械中关键的零部件,其运行状态直接影响到整个旋转机械的工作进程^[1-3]。因此,对滚动轴承进行更准确、更智能的故障诊断是减少经济损失的重要保障。

滚动轴承故障诊断包括数据采集、预处理、特征提取和故障分类等,其中特征提取和故障分类对诊断结果起到关键作用。而传统故障诊断技术需依赖于专家和技术人员对采集的数据进行多步骤繁琐的人工特征提取,无法满足“大数据时代”的要求^[4]。随着人工智能技术的发展,研究人员利用机器学习中支持向量机^[5](support vector machine, SVM)、BP神经网络^[6]和K近邻法^[7](K-nearest neighbor, KNN)等方法进行滚动轴承的故障诊断,虽然这些方法具有一定的非线性拟合能力,并在故障诊断领域取得一定的效果,但因其浅层的网络结构难以提取到深层的特征信息而降低滚动轴承故障诊断准确率。

深度学习是近年来机器学习发展中新型的研究热点,它由Hinton等人^[8]于2006年提出,因其强大的特征自动提取能力被学者应用到故障诊断领域。其中,Sun等^[9]提出一种基于稀疏自编码器的单层网络模型实现对感应电动机的故障分类;Li等^[10]基于深度置信网络(DBN)在滚动轴承故障诊断中取得较好的诊断精度。

深度学习中的卷积神经网络(Convolution neural network, CNN)是一种有监督的深度学习算法,它无需对原始采集到的故障数据进行预处理就能实现端到端滚动轴承故障诊断^[11-14]。张等^[15]提出基于自适应批量归一化(Adaptive batch normalization, AdaBN)算法的卷积神经网络模型,利用卷积神经网络自特征提取,实现了变工况下轴承故障诊断;雷亚国等^[16]提出了基于深度学习的机械装备大数据健康监测方法,能够自适应地提取故障数据中的特征信息,实现智能诊断;Lei等^[17]提出了一种基于长短时记忆网络(long short-term memory, LSTM)结合卷积神经网络的风力发电机故障诊断方法。虽然上述方法在各自的故障诊断任务中取得了较好的故障分类效果,故障诊断准确率相比传统诊断方法和机器学习方法高,但也受到以下两个因素的限制:其一,这些方法采用单尺度的卷积层提取特征,而滚动轴承在变

工况时,同种故障信息存在不同的频率变化,单尺度卷积网络不能有效提取到这些微小故障特征;其二,这些方法的网络结构较为复杂,未使用优化网络结构容易造成网络难以训练甚至退化等问题。

针对以上不足,本文提出了改进CNN的滚动轴承变工况故障诊断方法,设计了多尺度特征提取模块,能够有效挖掘故障数据中的特征信息;同时引入通道注意力机制帮助网络获得更重要的特征信息;设计了带跳跃连接线的卷积模块,能够获取更多前层网络的特征信息,使得故障诊断准确率更高。

1 基本方法

1.1 卷积神经网络

在深度学习中,卷积神经网络(CNN)是一种前馈神经网络,具有强大的特征自动提取能力,它通过构建多个卷积核将输入数据中的深层特征逐层提取出来,并且空间上降采样,达到降低输入维度的目的^[18]。CNN典型结构图如图1所示,主要由卷积层、池化层(下采样)和损失函数Softmax或SVM等分类器的全连接层构成。

卷积层主要是对输入数据进行特征提取,其内部存在多个卷积核,组成卷积核的每个元素包含一个权重系数和偏差。图1中的卷积操作可视为卷积核与前一层特征图(feature maps)做卷积,并通过激励函数进行非线性变换形成输出特征图,具体计算如式(1)所示:

$$X_{i,j}^{l+1} = f \left(\sum_{j=1}^L \sum_{i=1}^m (X_{i,j}^l \times w_{i,j}^l) + b \right) \quad (1)$$

式中,输入数据 $X_{i,j}^l$ 表示为网络第 l 层中第 i 个特征图的第 j 个特征值, L 表示为卷积核尺寸, $w_{i,j}^l$ 表示为权重系数, b 为偏差值, $f(\bullet)$ 为激活函数。

Softmax函数是Logistic分类器的推广,主要用于多分类问题。假设训练输入样本的集合为 $X = \{x_1 \dots x_i \dots x_T\}$,其中输入样本元素的类别为 $C = \{c_1 \dots c_i \dots c_K\}$ 集合中的一类,对应的标签为 \bar{y} ,于是将输入样本 x_i 判定为类别 C 集合中一类的概率为 $P(y_i = c_k | x_i)$ 。Softmax函数的数学表达式为:

$$P(y_i = c_k | x_i) = e^{x_i^T c_k} \times \frac{1}{\sum_{k=1}^K e^{x_i^T c_k}} \quad (2)$$

式中, $e^{x_i^T c_k}$ 表示为类别 c_k 与整个 x_i 分类类别之间的

相关关系, $1/\sum_{k=1}^K e^{x_k^T c_k}$ 表示为归一化函数。

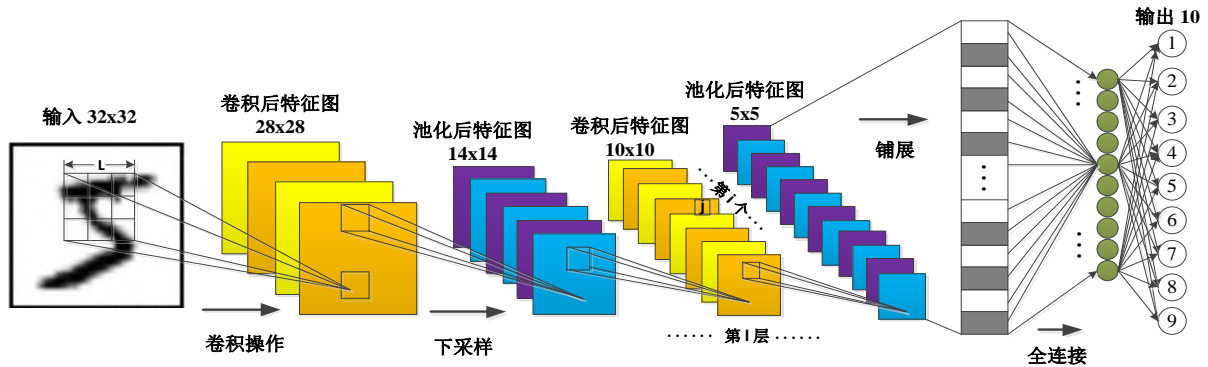


图1 卷积神经网络结构图

Fig.1 Structure of convolutional neural network

1.2 Inception 网络结构

Inception 模块是 GoogLetNet 神经框架中首次使用的优化模块^[19], 主要思想是利用不同大小的卷积核通过堆叠的方式增加网络宽度, 使其能够提取到丰富的特征信息, 同时利用 1x1 尺度的卷积核, 对输入的特征图进行降维, 减少参数计算量, 加快网络训练速度。目前, 经过几个版本的迭代, Inception 模块已经更新到 Inception-v4 和 Inception-ResNet 模块^[19]。其中, Inception-ResNet 模块通过结合 ResNet 中残差思想, 使得网络层数更深。

1.3 注意力机制

注意力机制最初被成功应用于机器翻译领域^[20], 现在已成为深度学习领域中一个重要概念, 并且是神经网络结构的重要组成部分。注意力机制通过借鉴人类的注意力方式被广泛应用在自然语言处理、统计学习及语音识别等领域中。

2 改进 CNN 的滚动轴承变工况故障诊断方法

滚动轴承一般工作在强噪声、变负荷等复杂的环境中, 采集的振动信号容易受到外界因素影响。为了充分发挥 CNN 网络特征提取能力, 同时预防梯度消失和退化问题, 本文提出了改进 CNN 的滚动轴承变工况故障诊断方法, 设计了多尺度特征提取模块和跳跃连接卷积模块并引入通道注意力机制, 实现对强噪声和变工况下故障数据的分类。

2.1 多尺度特征提取模块

为了从输入数据中最大化提取到特征信息, 本文设计了一种多层多通道的多尺度特征提取模块。该模块的设计主要基于 Inception 模块, 相比于传统

CNN 网络输入数据池化层, 它具有较强的特征提取能力, 其结构如图 2 所示: ①在第一层分别用 1x1、3x3 和 5x5 并行的卷积层对输入故障数据进行不同尺度的特征提取, 通道数设置为 16、8、8; ②第二层使用 5x5 和 3x3 的卷积层, 并串联在 3x3 和 5x5 的卷积层后面, 通道数都为 16; ③第三层使用 2 个 1x1 卷积层, 通道数为 32 和 16。为了优化故障数据, 提升网络诊断效果, 在每个卷积层后面都使用批量标准化和激活函数, 之后通过 Concat 层将不同分支的特征维度堆叠拼接到一起, 并使用通道注意力机制获得不同特征信息的重要程度, 依据重要程度增强有用特征, 抑制干扰特征, 最后通过最大池化层输出提取到的特征信息。

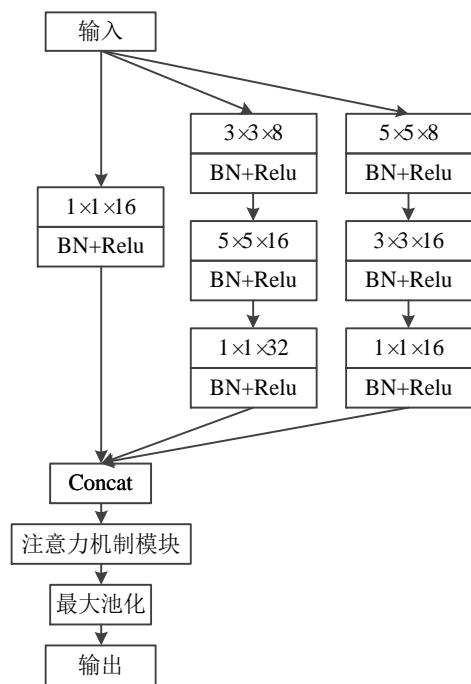


图2 多尺度特征提取模块结构

Fig.2 Structure diagram of multi-scale feature extraction module

2.2 带跳跃连接线的卷积模块

在传统卷积神经网络中，前一层的卷积层模块与后一层的卷积层进行端对端的首尾相连，不能有效利用该层输入向量与下一层之间的相关性，从而限制卷积网络对特征信息的学习效率。为此本文设计了带跳跃连接线的卷积模块，结构如图 3 所示。

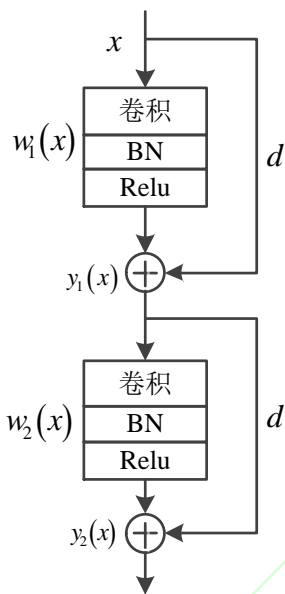


图 3 带跳跃连接线的卷积模块结构

Fig.3 Structure diagram of the convolution module with skip connection lines

图 3 中，在第一个卷积层输入向量和第二个卷积层之间添加了一条 d 倍的跳跃连接线，这样第二个卷积层不仅可以学习到第一个卷积操作后的特征信息，而且学习了前一层输入向量的传递信息，具有更高的学习效率。

假设跳线连接的卷积模块的输入为 x ，经过第一

个卷积层输出为 $y_1(x)$ ，第二个卷积层输出为 $y_2(x)$ ，则：

$$y_1(x) = w_1(x) + dx \tag{3}$$

$$y_2(x) = w_2(y_1) + dy_1 = w_2(w_1(x) + dx) + dw_1(x) + d^2x \tag{4}$$

由 $y_2(x)$ 输出可知，(4) 式含有第一层卷积模块输入向量 x 和 $dw_1(x)$ ，这使得下一个卷积模块接受到的特征信息增多，有利于将数据中微小故障特征提取出来，提高故障诊断的准确率。虽然跳跃连接能使每个卷积模块接受到更多的特征信息，然而这也导致整个网络模型最后的输出向量很庞大，参数的计算量增加严重，不利于滚动轴承的故障快速诊断。因此，在 Inception 模块之后，我们只加两层带跳跃线连接的卷积模块，从而保证整个网络的轻量化。

2.3 网络模型框架

本文提出改进 CNN 神经网络模型的结构图如图 4 所示，它以重构二维数据作为输入，然后进入多尺度特征提取模块，该模块通过不同尺度的卷积核和通道注意力机制获得更多更有效的故障数据信息，整个多尺度特征提取模块相当于一个卷积块，第二个卷积块为 Inception 模块，它由大小为 1、3、5 的卷积层和大小为 3 的最大池化层构成，通过对这些卷积层和最大池化层拼接，进一步提取故障数据中的特征信息。第三个卷积块则为带跳跃连接线的卷积模块，主要作用是增加网络通道数，学习更多特征信息。最后，通过全连接层把前三个卷积块提取的特征转化成一维数据，经过交叉熵损失函数输出故障诊断结果。

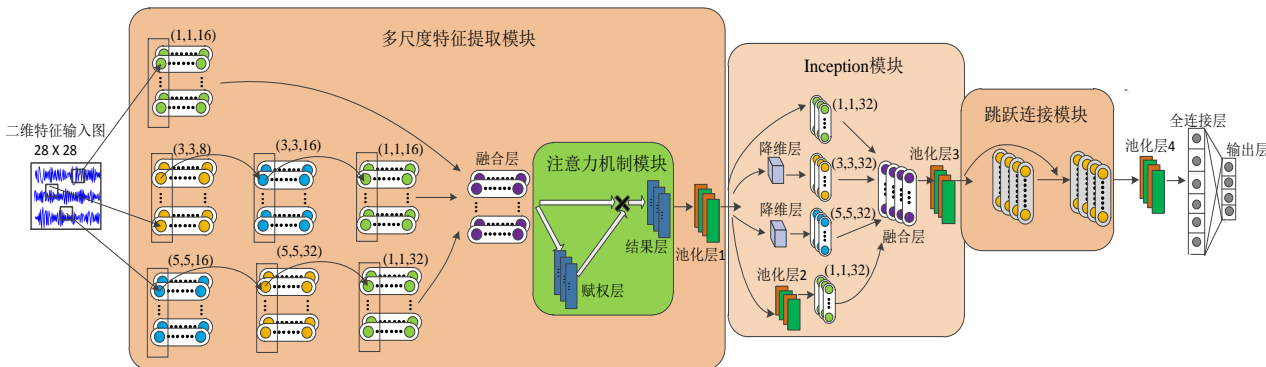


图 4 改进的 CNN 神经网络结构

Fig.4 Improved CNN neural network structure diagram

3 实验验证与分析

为了评估本文提出的改进 CNN 网络模型对故障诊断的有效性和准确性, 本文实验以滚动轴承为对象, 采用来自美国凯斯西储大学 (CWRU) 的轴承数据^[21]和东南大学变速箱数据集中的轴承数据进行实验验证^[22]。

3.1 CWRU 轴承故障数据集

3.1.1 故障数据集描述与处理 该数据被广泛用于滚动轴承故障诊断。CWRU 实验台由电机、扭矩传感器和测功器组成, 被测试的故障轴承是驱动端轴承型号为 SKF6205 电机。采用电火花方法分别在内圈、滚动体和外圈的表面上依次加工出损伤直径为 0.1778mm、0.3556mm、0.5334mm 和 0.7112mm 的单点凹槽, 以模拟滚动轴承在实际运行中出现的磨损。本实验采样频率为 12kHz, 分别采集了转速为 1797r/min、1772r/min、1750r/min 和 1730r/min 时, 对应负载为 0hp、1hp、2hp 和 3hp 状态下的加速度数据集, 并将其标记为数据集 A、数据集 B、数据集 C 和数据集 D。采集到的数据集根据不同位置以及损伤程度的不同划分为 16 种状态标签, 并且每种状态标签中的样本数量大致相同。采集的数据按照 3:1 的比例划分训练样本和测试样本, 每段的采样点设置为 784 点数。实验变工况数据集描述见表 1。

表 1 变工况实验数据集

Tab.1 Variable working condition experimental dataset

| 训练数据集 | 训练样本数 | 测试数据集 | 测试样本数 | 加噪声测试样本数 |
|-------|-------|-------|-------|----------|
| 数据集 A | 1972 | 数据集 B | 740 | 663 |
| | | 数据集 C | 740 | |
| | | 数据集 D | 740 | |
| 数据集 B | 2200 | 数据集 A | 663 | 740 |
| | | 数据集 B | 740 | |
| | | 数据集 D | 740 | |
| 数据集 C | 2200 | 数据集 A | 663 | 740 |
| | | 数据集 B | 740 | |
| | | 数据集 D | 740 | |
| 数据集 D | 2200 | 数据集 A | 663 | 740 |
| | | 数据集 B | 740 | |
| | | 数据集 C | 740 | |

(1) 重构输入数据格式。实验台采集的数据集为一维时间序列, 为了适应改进 CNN 网络模型输入数据格式, 有效地进行卷积和下采样操作, 本文将采样点数为 784 的一维数据重构为 [28,28] 的二维输

入特征图, 具体操作如图 5 所示, 将长度为 784 点数的一维时间序列进行等长截取, 每段有 28 数据点, 共 28 段, 最终获得 [28,28] 的二维输入特征图。

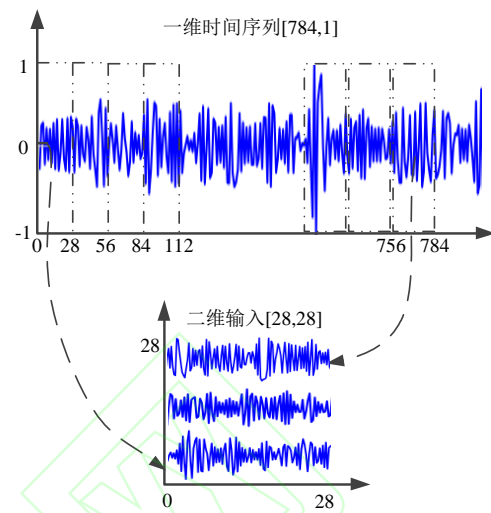


图 5 数据重构结构

Fig.5 Data reconstruction structure diagram

(2) 对输入数据进行数据标准化操作。为了更好地加快网络模型训练速度, 让数据便于计算和获得更加泛化的结果, 所以对输入数据做标准化处理, 数学表达式为:

$$X = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (5)$$

为了进行变噪声实验, 选取不同数据集下测试样本作为变噪声测试样本, 并且在测试样本中加入不同信噪比的高斯白噪声 (Signal to Noise Ratio, SNR) 以获取变噪声测试样本。

(3) 对诊断结果输出。为了评估归一化后的预测结果与对应样本标签之间的差异情况, 采用交叉熵损失函数进行误差损失值的计算, 相比于均方误差损失函数, 交叉熵损失函数更适应于目标输出为位矢量的分类问题, 数学表达式为:

$$J = -\sum_{i=1}^K I\{\bar{y}_i = k\} \log \frac{e^{x_i \cdot c_k}}{\sum_{k=1}^K e^{x_i \cdot c_k}} \quad (6)$$

式中, i 表示为第 i 个训练样本, k 表示为属于第 k 个类别, $I\{\cdot\}$ 表示为逻辑指示函数 (当值为真时, $I = 1$, 否则 $I = 0$), \bar{y}_i 表示为第 i 个真实标签。在训练的过程中不断调整参数, 达到最小化损失函数。

3.1.2 模型结构参数 实验是在 Windows10 系统、AMD 锐龙 5-4600H 处理器、16GB 内存的计算机上, 利用 pycharm 平台, 编程语言为 python, 使用

tensorflow 深度学习框架实现的。

网络模型的结构越深,对特征的表达能力越强,但是随着网络层数的增加,梯度爆炸以及退化问题也会表现出来,如滚动轴承的故障数据为一维时间序列,在转化成二维特征图像输入时,有效特征不如纯图像有效特征多,所以设计诊断模型时考虑网

络的轻量化。在实验过程中,采用 Adam 自适应优化器不断更新网络训练参数,批次大小设置为 64,迭代批数为 2200;使用动态学习率训练网络,学习步长设置为 0.001,衰减率为 0.9。该模型是在卷积神经网络的模型上构建的,因此参数设计与卷积神经网络类似,参数设计如表 2 所示。

表 2 参数设计

Tab. 2 Parameter design

| 层数 | 结构名称 | 结构参数 | 通道数量 | 输出大小 |
|----|----------------|-------------------------|------|-------|
| | 输入 | (28,28) | 1 | 28×28 |
| | (多尺度采样模块) | | | |
| 1 | 卷积层 | (1,1,1) | 16 | 28×28 |
| | 卷积层 | (3,3,1)/(5,5,1)/(1,1,1) | 32 | 28×28 |
| | 卷积层 | (5,5,1)/(3,3,1)/(1,1,1) | 16 | 28×28 |
| 2 | 通道注意力模块 | | | |
| 3 | 池化层 | (3,3,2) | | 14×14 |
| | (Inception 模块) | | | |
| 4 | 卷积层 | (1,1,2) | 32 | 7×7 |
| | 卷积层 | (1,1,1)/(3,3,2) | 32 | 7×7 |
| | 卷积层 | (1,1,1)/(5,5,2) | 32 | 7×7 |
| | 池化层 | (3,3,2)/(1,1,1) | 32 | 7×7 |
| 5 | 池化层 | (3,3,2) | | 4×4 |
| 6 | 跳跃连接模块 | (3,3,1) | 128 | 4×4 |
| | | (3,3,1) | 256 | 4×4 |
| 7 | 池化层 | (3,3,2) | | 2×2 |
| 8 | 全连接层 | (256/1024) | | 2×2 |
| 9 | Softmax 输出层 | | | |
| | 全连接层 | (1024/16) | | 2×2 |

表 2 中多尺度采样模块共有三个并行通道,分别从输入二维特征图中进行特征提取,如卷积层中参数(3,3,1)/(5,5,1)/(1,1,1)表示为:首先使用 3×3 卷积核进行特征提取,步长设置为 1;其次串接尺寸为 5×5 卷积核,步长为 1;最后使用大小为 1×1 卷积核,步长为 1。这样设计主要是减少计算参数,提升网络训练速度。Inception 模块中卷积层(1,1,1)/(3,3,2)也是同样的设计方法,不过卷积核步长为 2。

为了验证本方法在变噪声、变负荷和变工况实验中,是否能够获得较高的故障诊断准确率和较好的泛化能力,将本文实验结果与赵^[23]等人^[23]和朱^[24]等人提出方法的实验结果进行对比分析,其中,赵等人设计了 Inception + Residual block(简称 IRB)的方法,朱等人设计了 Attentional mechanism + Inception + Convolution neural network(简称 AICNN)的方法,方法的网络结构如表 3 所示。

表 3 不同方法的网络结构参数

Tab.3 Network structure parameters of different methods

| 方法 | 网络结构 |
|---------|---|
| Alexnet | C1(5,32)→P1(3,2)→C2(5,64)→P2(3,2)→C3(5,128)→C4(3,256)→P3(3,2)→FC1(256/1024)→FC2(1024/16) |
| IRB | 数据池化(3,3,2)→残差块 1(3,64)→残差块 2(3,64)→残差块 3(3,64)→残差块 4(3,128)→残差块 5(3,64)→FC1(256/1024)→FC2(1024/16) |
| AICNN | C1(3,32)→C2(3,64)→P1(3,2)→注意力机制→Inception 结构(128)→P2(3,2)→FC1(128/1024)→FC2(1024/16) |

3.1.3 变噪声故障诊断结果与分析 在实际工况中,滚动轴承不仅工作的环境里存在噪声,而且因自身零件震动和相互磨损也产生噪声,这些因素往往会

对采集到的振动数据产生污染, 所以进行变噪声实验在滚动轴承故障诊断中必不可少。本文以数据集 B 下的滚动轴承数据作为训练样本, 并且在该测试样本中加入信噪比分别为 3dB、6dB、9dB 和 12dB 的高斯白噪声, 最终将得到的实验结果与 IRB 和 AICNN 方法的诊断结果进行对比分析, 实验结果如表 4 所示。

表 4 不同方法的故障诊断准确率

| 故障诊断方法 | 信噪比/dB | | | |
|---------|--------|--------|--------|--------|
| | 3 | 6 | 9 | 12 |
| Alexnet | 76.89% | 90.54% | 94.19% | 95.67% |
| IRB | 79.86% | 92.29% | 95.00% | 95.67% |
| AICNN | 80.13% | 93.24% | 96.35% | 96.48% |
| 本文方法 | 88.17% | 97.70% | 98.32% | 99.19% |

由表 4 的对比数据可以看到, 在不同信噪比的实验结果中, 本文方法诊断的准确率总是优于 Alexnet、IRB 和 AICNN 的方法。基于 Alexnet 的方法因为缺乏跳跃连接卷积模块, 导致特征信息在前向传递时被丢失, 所以抗噪性能有一定下降; 基于 IRB 的方法, 虽然改进了数据池化层而增强了网络特征学习能力, 但是没有引入注意力机制并且网络结构过多使用残差块, 限制诊断的准确率; 基于 AICNN 的方法在输入数据池化层采用单尺度卷积核, 对输入二维图像数据特征提取能力不足, 当信噪比为 6dB 时, 故障诊断准确率为 93.24%, 相比于前两种方法准确率 90.56% 和 92.29% 提升不是很高; 本文方法在数据池化层采用多尺度卷积模块最大限度的提取了故障数据中的信息, 在 6dB 的噪声下故障诊断率达到 97.70%, 相比于前三种诊断方法其抗噪性能更强。

图 6 是本文方法在不同信噪比下的故障诊断结果图, 可以直观得到, 在 3dB 的强噪声环境下, 测试集的诊断准确率达到 88% 以上。

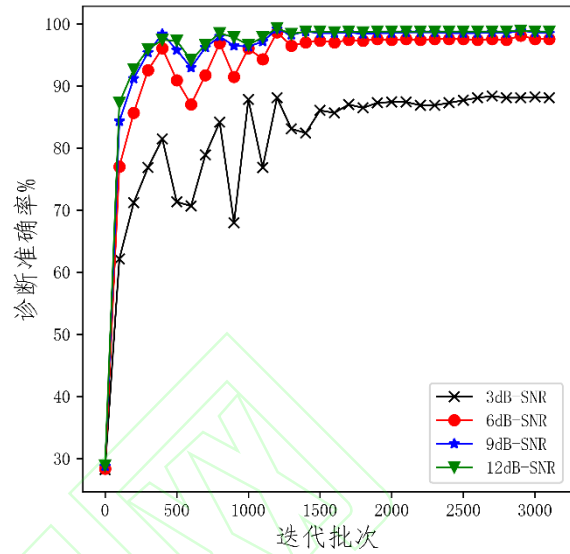


图 6 不同信噪比的结果

Fig.6 Results of different signal-to-noise ratios

3.1.4 变负荷故障诊断结果与分析 在实际工作中, 滚动轴承经常工作在变负荷状态, 所以要求故障诊断方法具有良好的泛化能力。为了进一步验证所提方法的泛化能力, 采用变噪声实验中的对比方法进行实验结果比较分析。为了模拟负荷变化情况, 以数据集 A、数据集 B、数据集 C 和数据集 D 中的一种作为训练样本, 另外三种负荷数据集作为测试样本。如图 7 中 A-B、A-C、A-D 表示为以数据集 A 作为训练样本, 分别以数据集 B、数据集 C 和数据集 D 作为测试样本, 变负荷实验结果如图 7 所示。

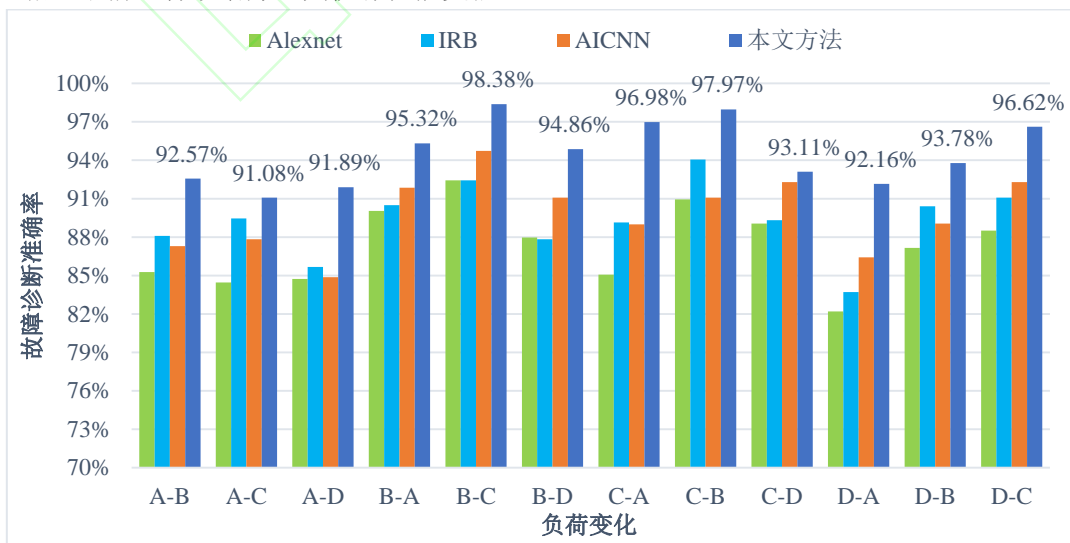


图 7 变负荷故障诊断结果

Fig.7 Diagram of fault diagnosis results of variable load

由图 7 可以看到, 本文提出的方法在变负荷实验结果中均高于其它三种方法, 其中以数据集 D 作为训练样本, 数据集 A、数据集 B 和数据集 C 作为测试样本时, 故障诊断准确率分别为 92.16%、93.78% 和 96.62%, 平均诊断率为 94.19%, 而基于 AICNN 方法的诊断平均值为 89.26%, 原因为本文方法采用了跳跃连接和多尺度特征提取, 能够充分提取故障数据中的信息, 从而使故障诊断准确率得到显著提升。

3.1.5 变工况故障诊断结果与分析 在变噪声和变负荷的单工况故障诊断实验中, 本文提出的方法均取得了显著的诊断效果, 接下来进一步研究在变工况情况下, 本文方法的诊断效果。具体过程为: 以数据集 B 作为训练样本, 以数据集 A、数据集 B、数据集 C 和数据集 D 作为测试样本并且向测试样本中分别添加 6dB、9dB 和 12dB 的高斯白噪声, 将取得的实验结果与 IRB 和 AICNN 比较, 结果如表 5 所示。

表 5 变工况故障诊断对比实验结果

Tab.5 Comparison experiment results of fault diagnosis under variable working conditions

| 信噪比/dB | 故障诊断方法 | 数据集 A | 数据集 B | 数据集 C | 数据集 D |
|--------|--------|--------|--------|--------|--------|
| 6 | IRB | 79.49% | 93.24% | 85.81% | 82.57% |
| | AICNN | 83.06% | 95.67% | 90.27% | 84.18% |
| | 本文方法 | 84.61% | 97.70% | 91.08% | 90.40% |
| 9 | IRB | 83.40% | 95.81% | 91.21% | 86.89% |
| | AICNN | 87.78% | 97.29% | 92.56% | 86.75% |
| | 本文方法 | 87.93% | 98.64% | 94.18% | 92.83% |
| 12 | IRB | 85.67% | 96.89% | 93.37% | 87.16% |
| | AICNN | 88.68% | 97.97% | 93.24% | 88.10% |
| | 本文方法 | 90.64% | 98.91% | 94.45% | 93.91% |

由表 5 可以得到, 本文所提方法在不同信噪比的噪声环境下, 故障数据诊断准确率都是优于其它两种比较方法。当信噪比为 6dB 时, 不同负荷测试样本的准确率平均值为 93.40%; 以负荷 3hp 为测试样本时, 随着信噪比的减少, 基于 AICNN 的诊断精度从 88.10% 降到了 84.18%, 而本文方法从 93.91% 降到了 90.40%, 诊断精度仍在 90% 以上。

为了检测所提方法的泛化能力, 本文以不同的

数据集作为训练数据并且加入 6dB 的高斯白噪声, 实验结果如图 8 直方图所示, 直方图的横轴表示负荷变化, 如 A-B 表示以数据集 A 负荷为训练样本, 数据集 B 作为测试样本, 纵轴表示测试准确率。由图 8 可知, 本文方法的柱高整体高于 IRB 和 AICNN 的柱高, 并且柱高更加平稳, 这说明本文方法诊断准确率更高, 泛化能力更强。

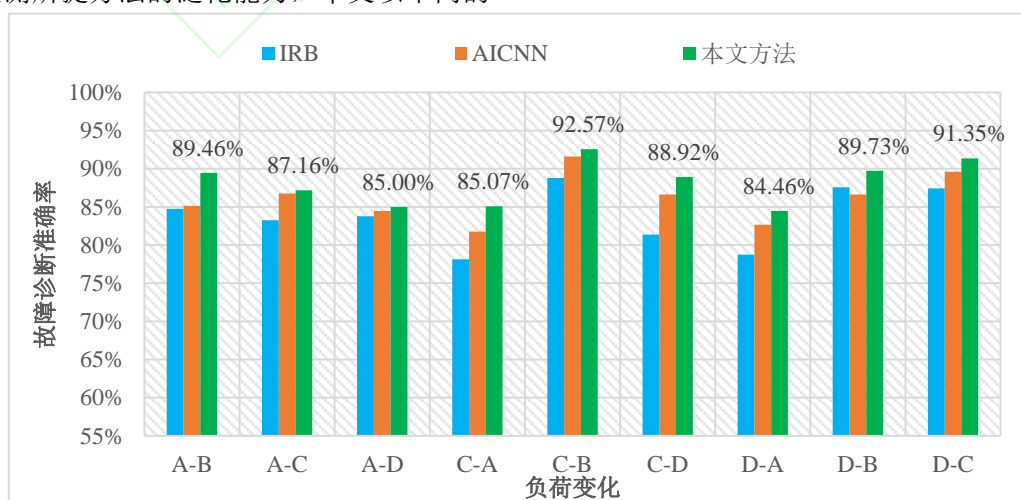


图 8 变工况诊断结果

Fig.8 Diagnosis results of variable conditions

3.2 变速箱数据集

3.2.1 故障数据集描述与处理 为了进一步验证所提方法的有效性, 本文采用来自东南大学变速箱数据集的轴承数据进行实验验证^[22], 该数据从传动系统动力学模拟器 (DDS) 上获取, 在转速-负载设置为 20Hz-0V 和 30Hz-2V 两种工况下, 共收集 8 个通道的数据。本实验采用通道 2 的数据, 它包括正常、滚动体故障、内圈故障、外圈故障和内外圈联合故障, 每种故障选取 600 个样本, 按照 3:1 的比例划分训练集和测试集, 整个数据集描述如表 6 所示。

表 6 东南大学轴承数据集划分

Tab.6 Classification of Bearing Data Set of Southeast

| 故障类型 | University | | | |
|--------|------------|-----|---------|-----|
| | 20Hz-0V | | 30Hz-2V | |
| | 训练集 | 测试集 | 训练集 | 测试集 |
| 正常状态 | 450 | 150 | 450 | 150 |
| 滚动体故障 | 450 | 150 | 450 | 150 |
| 内圈故障 | 450 | 150 | 450 | 150 |
| 外圈故障 | 450 | 150 | 450 | 150 |
| 内外联合故障 | 450 | 150 | 450 | 150 |

3.2.2 故障诊断结果与分析 该部分实验的模型结构与采用美国凯斯西储大学轴承数据集类似, 具体模型结构参数见表 2 所示。

在转速-负载分别设置为 20Hz-0V 和 30Hz-2V 两种工况下, 以轴承 5 种故障类型为基础, 构建 10 个实验数据集。为了避免实验结果的偶然性, 对 5 次实验结果求平均值, 同时将实验结果与 Alexnet、IRB 和 AICNN 的方法作比较, 结果如图 9 所示。

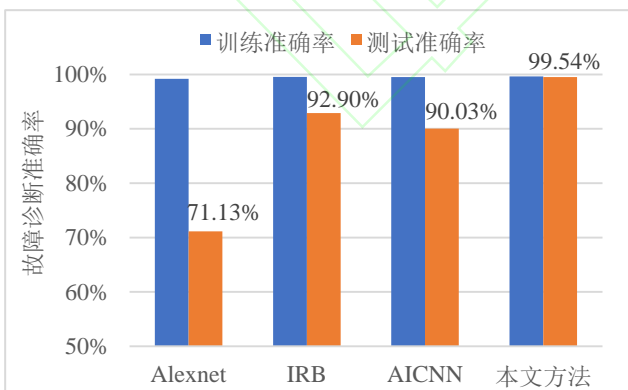


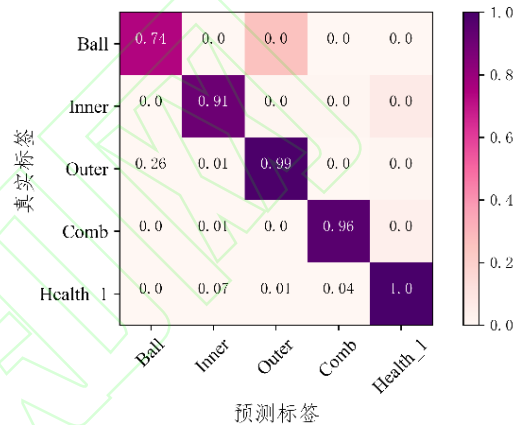
图 9 不同方法故障诊断结果

Fig.9 Diagram of fault diagnosis results of different methods

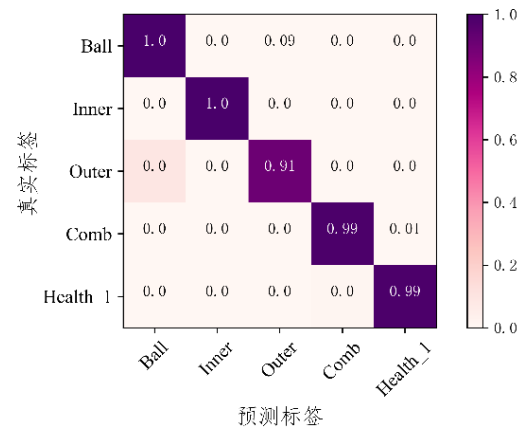
由图 9 可知, 本文所提方法的故障诊断准确率相比较其它三种方法具有显著优势, 其诊断准确率达到 99.54%, 原因是本文方法设计了跳跃连接模块

和多尺度特征提取模块, 能够充分提取故障数据中的信息, 从而使故障诊断准确率得到显著提升。

为了更进一步考察本文所提算法对故障误判的情况, 对测试结果做了混淆矩阵实验, 其中, x 轴横坐标为预测标签的诊断状态, y 轴纵坐标为真实标签的实际状态。从图 10 (a) 中可知, IRB 方法对于轴承滚动体故障和内圈故障的诊断准确率仅仅只有 74% 和 91%, 而从图 10 (b) 中可知, 本文所提方法除了将 9% 的 Ball 故障误诊为 Outer 故障外, 其它状态均有较高的诊断准确率, 其整体诊断准确率达 97.80%。这表明本文所提方法在不同的实验数据集上, 仍然具有较高的诊断准确率并且具有更好的泛化能力。



(a) IRB 方法



(b) 本文方法

图 10 不同方法的混淆矩阵图

Fig.10 Confusion matrix diagram of different methods

4 结论

本文提出了一种改进 CNN 的滚动轴承变工况故障诊断模型, 它是基于 Inception 网络结构设计数据池化层并引入注意力机制搭建而成, 能够解决滚动轴承运行时负荷变化和强噪声干扰导致故障诊断率低和泛化能力差的问题。主要结论如下:

(1) 设计了多尺度特征提取模块, 通过多通道不同尺度卷积层从输入数据中进行特征提取, 从而保证最大限度地提取到故障数据中的有效信息; 在多尺度特征提取模块中的融合卷积层之后, 引入通道注意力机制模块, 以此实现自适应地增强有效信息、抑制干扰信息; 在全连接层之前, 设计带跳跃连接线的卷积模块, 增强卷积模块对前一模块的特征学习效率, 减少信息丢失; 整个模型框架相当于由3个卷积块组成, 易于训练, 满足轻量化要求。

(2) 由实验验证可知, 与几种深度网络模型相比, 本模型不仅在强噪声和变负荷的环境里有着极高诊断准确率, 而且在变工况环境中也有较高的准确率。

(3) 由变工况实验可知, 本文的方法在不同训练集下, 不同测试集的故障诊断准确率更高, 这表明本网络的泛化性优于其它几种相比较的网络。

参考文献

- [1] WANG Huaqing, KE Yanliang, LUO Ganggang, et al. Compressed sensing of roller bearing fault based on multiple down-sampling strategy [J]. *Measurement Science and Technology*, 2016, 27(2): 025009.
- [2] JIANG Hongkai, XIA Yong, WANG Xiaodong. Rolling bearing fault detection using an adaptive lifting multiwavelet packet with a $1\frac{1}{2}$ dimension spectrum [J]. *Measurement Science and Technology*, 2013, 24(12): 125002.
- [3] 张妮, 车立志, 吴小进. 基于数据驱动的故障诊断技术研究现状及展望 [J]. *计算机科学*, 2017, 44(Z1): 37-42.
ZHANG Ni, CHE Lizhi, WU Xiaojin. Research status and prospects of data-driven fault diagnosis technology[J]. *Computer Science*, 2017,44(S1):37-42.
- [4] 张西宁, 郭清林, 刘书语. 深度学习技术及其故障诊断应用分析与展望 [J]. *西安交通大学学报*, 2020, 54(12): 1-13.
ZHANG Xining, GUO Qinglin, LIU Shuyu. Analysis and prospect of deep learning technology and its application in fault diagnosis[J]. *Journal of Xi'an Jiaotong University*, 2020,54(12):1-13.
- [5] 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: 436-446.
- [6] GU Yingkui, ZHOU Xiaoqing, YU Dongping, et al. Fault diagnosis method of rolling bearing using principal component analysis and support vector machine [J]. *Journal of Mechanical Science and Technology*, 2018, 32(11): 5079-5088.
- [7] 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, LU Feiya, BAO Zhejing, et al. Review of small fault diagnosis methods based on data-driven[J]. *Acta Automatica Sinica*, 2016,42(9):1285-1299.
- [8] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786): 504-507.
- [9] SUN Wenjun, SHAO Siyu, ZHAO Rui, et al. A sparse auto-encoder-based deep neural network approach for induction motor faults classification [J]. *Measurement*, 2016, 89: 171-178.
- [10] LI Chuan, SANCHEZ R V, ZURITA G, et al. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis [J]. *Neurocomputing*, 2015, 168: 119-127.
- [11] JIANG Guoqian, HE Haibo, YAN Jun, et al. Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox [J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(4): 3196-3207.
- [12] JIA Feng, LEI Yaguo, LU Na, et al. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization [J]. *Mechanical Systems and Signal Processing*, 2018, 110: 349-367.
- [13] WANG Shuhui, XIANG Jiawei. A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps [J]. *Soft Computing*, 2020, 24(4): 2983-2997.
- [14] 宫文峰, 陈辉, 张泽辉, 等. 基于改进卷积神经网络的滚动轴承智能故障诊断研究 [J]. *振动工程学报*, 2020, 33(2): 400-413.
GONG Wenfeng, CHEN Hui, ZHANG Zehui, et al. Intelligent fault diagnosis of rolling bearing based on improved convolutional neural network [J]. *Journal of Vibration Engineering*, 2020, 33(2): 400-413.
- [15] ZHANG Wei, PENG Gaoliang, LI Chuanhao, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals [J].

- Sensors, 2017, 17(2): 425.
- [16] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法 [J]. 机械工程学报, 2015, 51(21): 49-56.
- LEI Yaguo, JIA Feng, ZHOU Xin, et al. Mechanical equipment health monitoring method based on deep learning theory [J]. Journal of Mechanical Engineering, 2015,51(21):49-56.
- [17] LEI Jinhao, LIU Chao, JIANG Dongxiang. Fault diagnosis of wind turbine based on long short-term memory networks [J]. Renewable Energy, 2019, 133: 422-432.
- [18] 宫文峰, 陈辉, 张美玲, 等. 基于深度学习的电机轴承微小故障智能诊断方法 [J]. 仪器仪表学报, 2020, 41(1): 195-205.
- GONG Wenfeng, CHEN Hui, ZHANG Meiling, et al. Intelligent diagnosis method for incipient fault of motor bearing based on deep learning[J]. Chinese Journal of Scientific Instrument, 2020, 41(1): 195-205.
- [19] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4: inception-resnet and the impact of residual connections on learning [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2017:4278-4284.
- [20] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [EB/OL]. [2021-04-01]. <https://arxiv.org/abs/1409.0473>
- [21] 袁壮, 董瑞, 张来斌, 等. 深度领域自适应及其在跨工况故障诊断中的应用 [J]. 振动与冲击, 2020, 39(12): 281-288.
- YUAN Zhuang, DONG Rui, ZHANG Laibin, et al. Deep domain adaptation and its application in fault diagnosis under different working conditions[J]. Vibration and shock, 2020,39 (12): 281-288
- [22] SHAO Siyu, MCALEER S, YAN Ruqiang, et al. Highly accurate machine fault diagnosis using deep transfer learning [J]. IEEE Transactions on Industrial Informatics, 2019, 15(4): 2446-2455.
- [23] 赵小强, 梁浩鹏. 使用改进残差神经网络的滚动轴承变工况故障诊断方法 [J]. 西安交通大学学报, 2020, 54(9): 23-31.
- ZHAO Xiaoqiang, LIANG Haopeng. Fault diagnosis method for rolling bearing under variable working conditions using improved residual neural network [J]. Journal of Xi'an Jiaotong University, 2020, 54(9): 23-31.
- [24] 朱浩, 宁芊, 雷印杰, 等. 基于注意力机制-Inception-CNN 模型的滚动轴承故障分类 [J]. 振动与冲击, 2020, 39(19): 84-93.
- ZHU Hao, NING Qian, LEI Yinjie, et al. Fault classification of rolling bearings based on attention mechanism-inception-CNN Model[J]. Journal of Vibration and Shock, 2020,39(19):84-93.

(编辑 陶晴)