DOI: 10.13382/j.jemi.B2105070

改进卷积 Lenet-5 神经网络的轴承故障诊断方法*

赵小强[1,2,3] 罗维兰[1,2]

(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:针对滚动轴承微弱信号在强噪声、变工况复杂环境下,难以实现有效的故障诊断问题,提出了一种改进卷积 Lenet-5 神经网络的轴承故障诊断方法。首先,对采集的一维时域轴承振动信号进行预处理转化成便于卷积操作的二维灰度图;其次,将最基本的 Lenet-5 模型中的连续单向的传统卷积层改进为 Block1 模块、Block2 模块、Block3 模块,提取到更完整、更精准的特征信息;最后,为了防止网络出现过拟合现象,采用 L2 正则化和 Dropout 优化网络。为了验证本文所提方法在复杂工况环境的鲁棒和泛化性能,利用滚动轴承数据集和变速箱实验数据集进行实验验证。轴承数据集实验结果表明,本文所提出的方法在变噪声实验中准确率平均值都在 99.3%;在变负荷实验中,故障诊断准确率都高于 90.26%;在变工况实验中,故障诊断准确率平均值都高于 89.01%;在变速箱数据集实验中,抗噪性故障诊断准确率高达 96.3%。采用改进的 Lenet-5 方法对滚动轴承 12 种故障类型具有更好的分辨能力,在变工况下具有更好的抗干扰性和泛化性能。

关键词:滚动轴承;故障诊断;复杂工况;Lenet-5 网络;网络优化

中图分类号:TH133.33;TN06 **文献标识码**:A **国家标准学科分类代码**:510.40

Improved convolutional Lenet-5 neural network for bearing fault diagnosis

Zhao Xiaoqiang[1,2,3] Luo Weilan[1,2]

(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 that it is difficult to realize effective fault diagnosis for weak signals of rolling bearings in the complex environment of strong noise and variable working conditions, a bearing fault diagnosis method based on improved convolutional Lenet-5 neural network is proposed. Firstly, the collected one-dimensional time-domain bearing vibration signals are preprocessed and converted into two-dimensional grayscale images which are convenient for convolution operation. Secondly, the continuous one-way traditional convolutional layers in the most basic Lenet-5 model are improved into Block1 module, Block2 module and Block3 module to extract more concrete and accurate feature information. Finally, L2 regularization and Dropout optimization are used to avoid overfitting. In order to verify the robustness and generalization performance of the proposed method in complex working conditions, experimental validation was carried out using the rolling bearing dataset and the gearbox experimental dataset. The experimental results of the bearing dataset show that the average accuracy of the proposed method in the variable noise experiments is 99.3%. In the variable load experiments, the average accuracy of fault diagnosis is higher than 90.26%. In the variable operating conditions experiments, the average accuracy of fault diagnosis is higher than 89.01%. In the gearbox dataset experiments, the fault diagnosis accuracy of anti-noise is up to 96.3%. The improved Lenet-5 method has the better ability of fault diagnosis for 12 fault types of rolling bearings, and has better anti-interference and generalization performance under variable working conditions.

Keywords: rolling bearing; fault diagnosis; complex working conditions; Lenet-5 network; network optimization

收稿日期:2021-12-30 Received Date: 2021-12-30

*基金项目:甘肃省科技计划(21YF5GA072)、国家重点研发计划(2020YFB1713600)、甘肃省教育厅产业支撑计划(2021CYZC-02)项目资助

0　引　言

随着工业化快速发展,旋转机械运行的工况复杂、工作环境恶劣,导致滚动轴承故障引起的冲击特征常常淹没在强噪声背景信号中,难以识别[1],这大大增加了故障发生的概率。而滚动轴承作为旋转机械设备关键的零部件,对旋转机械的性能、稳定性起到至关重要的作用,约1/3的旋转机械故障是因滚动轴承引起的[2-4],因此能够实现滚动轴承故障诊断具有重要的意义。

近几年来,深度学习被广泛应用到图像处理、计算机视觉等多个领域,在故障诊断中也逐渐被使用,利用深度学习实现复杂工况下端到端的轴承故障诊断[5]。深度学习中的卷积神经网络是近年来发展起来的一种高效特征识别和分类方法[6]。

曲建岭等[7]首先对数据进行重叠分段预处理,然后对预处理的样本采用一维卷积神经网络进行网络自适应特征提取,实现轴承故障诊断。叶壮等[8]采用经验模态分解[9](EMD)将一维时序信号分解处理为若干通道的二维特征图,有利于CNN卷积特征提取,对多通道所提取的特征信息基于峭度加权融合,最后实现故障分类。董绍江等[10]采用奇异值分解(SVD)[11]将数据处理成原始降噪信号和带噪声信号两部分,将带噪声信号经过SVD去噪和EMD分解进行特征提取,提取的特征输入到改进卷积神经网络(ICNN)训练,实现轴承故障诊断。徐彦伟等[12]采用小波包分解[13-14]对原始信号处理,再通过卷积神经网络训练,实现对不同工况下的地铁牵引电机轴承故障诊断。

以上方法表明CNN卷积在滚动轴承故障诊断的可行性,然而以上方法未考虑轴承通常运行在大型旋转机械中,振动信号常受到环境噪声和载荷变化的影响,因此,所采集的振动信号在多个时间尺度上的特征频率变化较大,具有高度的复杂性、耦合性和不确定性的特性。针对振动信号具有复杂特征,单尺度的卷积不能充分提取特征信息的问题,研究人员已经将CNN与多尺度特征结合起来,利用多尺度CNN学习振动信号的多尺度特征。许子非等[15]首先利用离散小波变换得到不同尺度下振动信号的系数,然后利用多个全卷积网络提取不同尺度下的特征,最终实现了风力机叶片结冰检测。Peng等[16]将传统的滤波方法与CNN相结合,提出了一种多分支、多尺度的CNN,能够更准确地诊断高速列车轴承故障。

以上方法尽管取得不错的效果,但是对于多尺度CNN仍忽略了一些问题:1)多尺度CNN卷积核分布相对单一,所提取的特征信息不能包含丰富的信号特征;2)由于噪声的干扰,特征提取的过程中可能因无法检测到故障相关的特征信息,而丢失一些关键的特征信息;3)多尺度CNN采用多个并行网络结构,网络结构复杂,网络训练困难。使得输出层学到的高级特征往往不足以正确地分类故障。

因此,本文基于最基本的模型前提下,提出了改进卷积Lenet-5神经网络的轴承故障诊断方法。在不依赖预处理复杂信号的前提,改进卷积后的Lenet-5具有强大的自学习能力,可以直接从原始信号中自适应、高效的端到端提取到更完整、更精准的深层特征信息,实现变工况下的轴承故障诊断。通过两组实验数据验证了本文所提方法的有效性,实验结果表明,本文所提的方法比其他方法更具有抗噪性和泛化能力。

1　基础理论

1.1　Lenet-5模型结构

Lenet-5模型[17]是一种最经典的网络结构,广泛地应用到手写数字识别、面目表情等领域,网络结构的参数根据所解决的问题的难易程度进行调整改进,模型的基本结构如图1所示。主要由输入层、特征提取、全连接层以及故障分类4部分构成,其中特征提取包括3层卷积层($C_i, i=1,2,3$)和2层池化层($S_i, i=1,2$)。

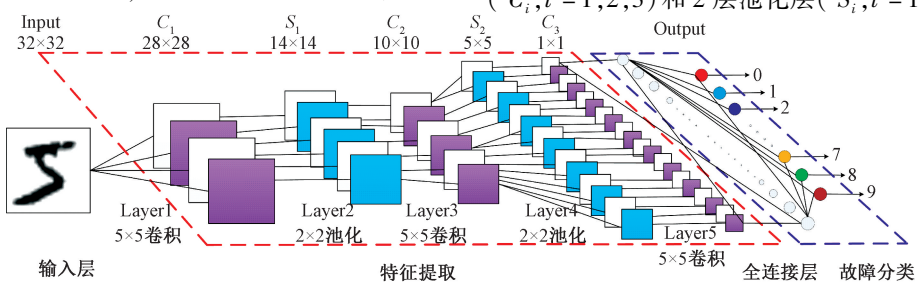


图1　Lenet-5模型结构图

Fig. 1　Lenet-5 model structure diagram

Lenet-5模型输入是32×32图像,经过3层卷积核为5×5的卷积操作,卷积操作主要是对输入的数据进行特征提取和特征映射,得到不同特征信息,卷积层运算如式(1)所示。

$$x_i^l = \sum_{j=0}^{W-1} k_{ij}^l * x_j^{l-1} + b_i^l \tag{1}$$

式中：k_{ij}^l 代表第 l 层的第 i 卷积核的第 j 个权值；x_j^{l-1} 代表第 $l-1$ 层第 j 个被卷积的局部区域，b_i^l 代表第 l 层的第 i 卷积核的偏置，W 代表卷积核的大小。为了减少网络数据运算量、降低空间维度以及增加网络泛化能力，通常在卷积层后面添加池化层，对卷积层传递的特征图进行特征选择和信息过滤，提高模型的容错性；池化窗口大小为 2×2，步长为 2，池化层计算如式(2)所示。

$$y^{l(i,j)} = \max_{(j-1)W+1 \leqslant t \leqslant jW} \{x^{l(i,t)}\} \tag{2}$$

式中：$y^{l(i,j)}$ 代表池化后区域大小，$x^{l(i,t)}$ 代表第 i 帧第 t 个神经元激活，W 代表池化区域大小。整个 Lenet-5 模型的具体参数如表 1 所示。

表 1　Lenet-5 模型具体参数

Table 1　Specific parameters of Lenet-5 model

结构名称	网络层名称	卷积核	步长	通道数	输出大小
输入	—	—	—	—	32×32×1
特征提取	卷积 Layer1	5×5	—	6	28×28×6
	池化 Layer2	2×2	2	—	14×14×6
	卷积 Layer3	5×5	—	16	10×10×16
	池化 Layer4	2×2	2	—	5×5×16
	卷积 Layer5	5×5	—	120	1×1×120
全连接层	全连接层	—	—	84	—
输出分类	—	—	—	10	—

1.2　深度残差收缩网络

深度残差收缩网络[18]是一种从强噪声振动信号中学习鉴别特征的潜在有效方法，实质上是深度残差网络、注意力机制和软阈值函数的集成，将软阈值作为可训练的收缩函数插入到深度结构中，以强制不重要的特征为零，从而使学习到的高级特征，有效地消除噪声相关特征。

深度残差网络框架图如图 2 所示，首先，网络进行两层卷积操作，其作用主要是将重要特征信息转化成绝对值较大的值，将冗余的特征信息转化为绝对值小的值；然后，通过虚线框中的虚线箭头将卷积后的特征取绝对值，全局平均池化和求平均，训练学习找到最大值与最小值的界限一组特征，记为 A，虚线框中实线箭头是全局均值池化后的特征连接到全连接网络中，经 Sigmoid 函数将输出值调整为 0~1 之间，记为 α；其次，输出的阈值为 $A\times\alpha$，这种阈值(相当于注意力机制)不仅为正数，而且阈值不会太大；最后，残差连接的恒等映射便于反向传播，从而更容易训练得到更优的参数。

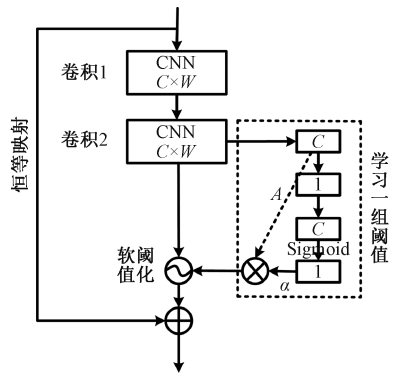


图 2　残差收缩网络

Fig. 2　Residual shrinkage network

2　改进卷积神经网络 Lenet-5 的轴承故障诊断

2.1　模型网络框架结构

实际工业环境中，滚动轴承常工作于强背景噪声、变转速负荷环境下运行，亟需一种简单、高效自适应端到端提取关键特征信息的方法，实现对强噪声、变负荷、变工况复杂环境的轴承故障诊断分类。

然而深度学习中的卷积操作过程中，由于传统 CNN 通常使用单个卷积核，所提取的特征信息不充分，不能全面地反映轴承的健康状态。随着多尺度 CNN 的出现有效地解决特征提取不充分的问题，但多尺度提取的重要特征会淹没在强噪声背景下，同时增加不必要的冗余信息，从而增加网络训练的参数，不利于神经网络模型学习时域信号的特征信息。因此，本文提出一种改进卷积 Lenet-5 神经网络的轴承故障诊断方法，其结构如图 3 所示，主要由信号处理、特征提取、全连接层和故障分类 4 部分构成。首先，将采集的一维时序滚动轴承数据集，通过横向插样[19]的方法预处理成便于卷积操作的二维特征图，轴承数据采集频率点为 1 024，按顺序依次插入成 32×32 的特征图；然后，采用改进卷积层 3 个 Block 模块代替传统的卷积层，提高卷积特征提取的能力，将改进的卷积层 Block 模块应用在最经典的 Lenet-5 模型中，验证改进卷积层特征提取的性能；其次，全连接层采用 L2 正则化和 Dropout 进行网络优化，快速训练到最优参数，提高网络特征学习能力；最后，Softmax 分类器实现滚动轴承故障分类。

2.2　改进卷积层 Block 模块设计

改进卷积层 Block 模块如图 4 所示。Block 模块结构中，首先，采用不同尺度的卷积核所提取的特征信息精度不同，卷积核尺度越小，提取的特征信息更能体现出局部信息，卷积核尺度越大，提取的特征信息更能体现出全局

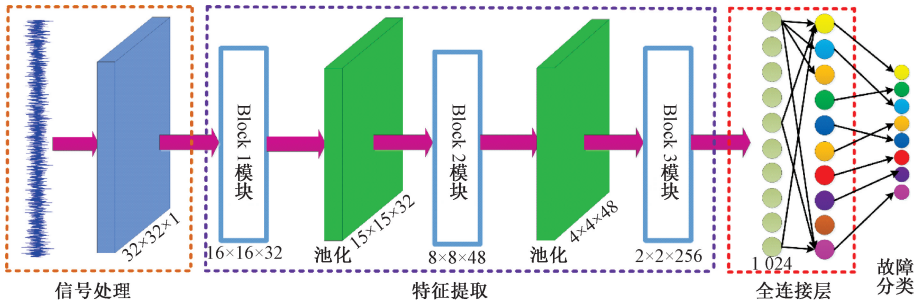


图3　改进的模型框架

Fig. 3　The improved model frame diagram

信息,多尺度卷积核对振动信号中挖掘出更深层、更丰富的多尺度特征信息,并将多尺度特征信息通过特征维度拼接融合;然后,将残差收缩网路中的软阈值作为可训练的收缩函数插入到深度结构中,以强制不重要的特征为零,从而使学习到的高级特征变得更有分辨力;最后,通过残差连接特征映射加快网络收敛,提高网络特征学习能力。

Block1 模块结构如图4(a)所示,首先,因 Block1 模块多尺度卷积是从原始信号中特征提取的第1步,所提取的特征信息直接影响着最终的故障分类,因此 Block1 模块中的多尺度提取特征信息要求:既包含丰富的全局信息,又包含丰富的局部信息;而传统单层且只使用同一尺度的卷积核,容易忽略不同时间尺度的原始信号中的长期和短期特征,导致提取的特征表达信息不够完整[20],因此本文设计的 Block1 模块首先采用不同尺度的小卷积核提取更加精细的特征信息,设计4个分支,上面两个支路浅层使用大卷积 5×5 和 3×3,提取全局信息,后面深层使用 1×1 的小卷积核提取深层抽象信息,中间支路与前面两支路特征提取所使用的卷积核大小相反,主要目的是与之互补,提取更完整的特征信息,最后一个分支采用 1×1 卷积核主要是数据降维减少参数量和提高泛化性能,这样4个支路所提取的特征信息通过特征维度拼接融合,可以得到更充分特征信息;其次,将融合的特征信息经过残差收缩网络进行软阈值化和注意力机制,根据数据集中噪声强度的不同,自适应学习一组不同的阈值,使得冗余信息的特征置为0,从而有效地提高对高噪声振动信号的鉴别特征学习能力,选择出更加有效的特征信息;最后,采用残差连接避免梯度爆炸,并加快网络的收敛性。

由于 Block1 模块采用多尺度特征信息既包括局部信息也包括全局信息,在提取特征中增加了网络的参数量,借鉴文献[21]采用局部稀疏网络结构的维度缩减版本提取特征信息可以降低网络的参数量并提高网络的训练速率的思想,本文 Block2 模块采用局部稀疏网络进行特征提取,如图4(b)所示,可以减少网络参数量并提高网络的训练效率;局部稀疏网络池化层可以有效避免在迭代过程中参数的调整,并保留边缘的关键信息;局部稀疏网络的每个小卷积层后都采用批量规范化(BN)和激活函数(ReLU),增强网络的学习能力;然后将局部稀疏网络提取的特征信息进行特征维度拼接融合,同样使用残差收缩网络有效地提高对高噪声振动信号的鉴别特征学习能力,并用残差网络特征映射加快网络收敛速度,防止网络出现过拟合现象。

Block3 模块如图4(c)所示,Block3 模块与 Block1 模块类似,同样对多尺度提取特征信息采用四个分支,与 Block1 模块不同的是:将 Block1 模块中为了提取全局特征信息的 5×5 大卷积核分支替换成两个 3×3 的小卷积核进行操作,为了提取到更关键、更精细的特征信息。

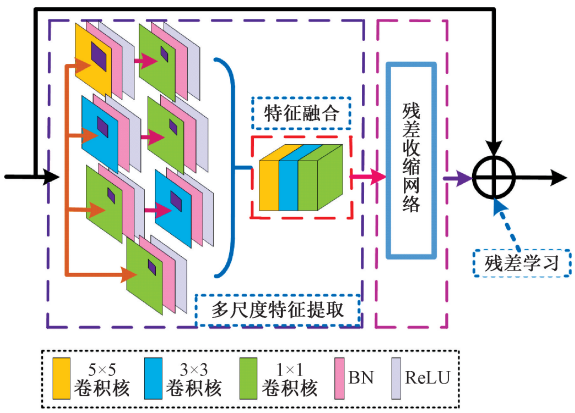
Block1 模块提取的特征既包含全局信息又包含局部信息,因此放入卷积层的第1层,浅层中提取到丰富振动信号;Block3 模块与 Block1 模块结构相似,但是多尺度的卷积核变小,希望能从深层中挖掘出高级特征;而 Block2 模块采用局部系数网络主要是将 Block1 模块提取的特征降维,降低网络参数量,同时提高了网络的训练效率。这样3个 Block 模块组合,使网络在强噪声、复杂工况环境下,具有更高的特征学习能力和诊断性能。

2.3　网络算法优化

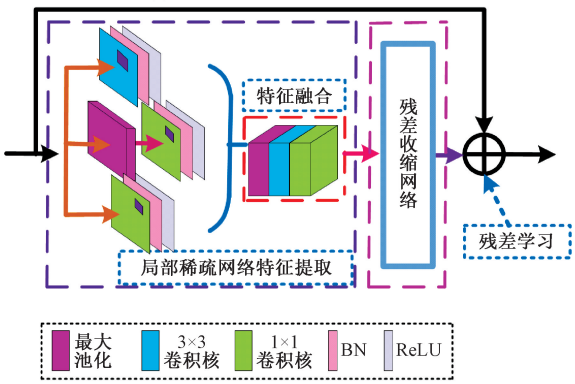
改进的网络相对基本网络更复杂,所提取的特征更有效、更完整,但是随着网络的复杂,容易造成参数量大,在训练集相对较小的情况下容易造成过拟合现象,因此采用 L2 正则化方法[22]和 Dropout 算法对网络进行优化。

1)L2 正则化算法

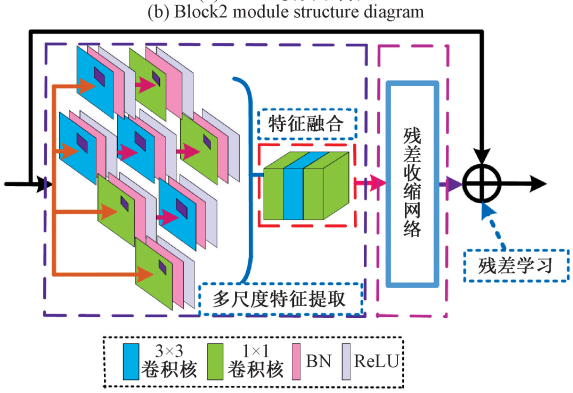
正则化是在机器学习过程中可以降低模型复杂度和不稳定程度且常用的一种技术,可以避免训练的结果过拟合化。L2 正则化原理是:在原来的损失函数式(3)基础上加入式(4)得到式(5),在训练过程中,加入 $R(w)$ 中拟合 w 可以抑制原损失函数拟合的 w 值;使得式(5)中 $J(w,b)$ 的 w 值变小,w 值越小,网络模型复杂度越低,过拟合程度越低。



(a) Block1模块结构图
(a) Block1 module structure diagram



(b) Block2模块结构图
(b) Block2 module structure diagram



(c) Block3模块结构图
(c) Block3 module structure diagram

图 4　改进卷积层 Block 模块结构

Fig. 4　The improved convolutional layer Block module structure diagram

$$J_0(w,b) = -\frac{1}{m}\sum_{i=1}^{m} L(\hat{y},y) = -\frac{1}{m}\sum_{i=1}^{m}(\hat{y} - (b + \sum w_i x_i))^2 \tag{3}$$

$$R(w) = \frac{\lambda}{2m}\| w \|_2^2 = \frac{\lambda}{2m}\sum_{i=1}^{m} w_i^2 \tag{4}$$

$$J(w,b) = -\frac{1}{m}\sum_{i=1}^{m} L(\hat{y},y) + \frac{\lambda}{2m}\| w \|_2^2 \tag{5}$$

式中：$\hat{y}$ 神经元输出的期望值，$\| w \|_2^2$ 为 w 的 L2 范数值，b 为神经元传递偏置量，λ 为正则化参数，m 为样本数据集的大小，本文经过交叉验证，λ 取值为 0.005。

2) Dropout 算法

Dropout 作为一种常用的优化算法，在网络训练过程中每次迭代时，神经元的输出值以特定概率“丢弃”一些节点，即让一部分神经元“失活”，从而防止网络训练出现过拟合现象[23]。保持的概率在 0~1 之间，本文采用交叉法验证 Dropout 参数值设定为 0.5。

3　实验数据与参数

本文仿真实验采用 Python 编程语言，Tensorflow 学习框架，硬件平台为 Intel(R) Core(TM) i5-4200 H CPU，内存：8.00 GB。

3.1　实验数据

实验 1：数据选自美国西储凯斯大学(Case Western Reserve University)轴承研究所[24]，该数据使用的实验装置如图 5 所示，实验装置主要包括电机、驱动端轴承、扭矩传感器及编码器、测力器。驱动端的滚动轴承型号为 SKF6205，利用电火花放电的方式在轴承的内圈、外圈、滚动体加工单点损伤，损伤直径尺寸大小分别为 0.177、0.355、0.533、0.711 mm，轴承振动的信号由加速度传感器采集，采集的频率为 12 kHz，电机的载荷分别为 0.476、1.492、2.238 kW，分别记作 Bhp、Chp、Dhp，另外空载负荷，记作 Ahp。

电机根据轴承不同位置所发生的损伤直径不同分 11 种故障状态和一种正常状态，构成 12 种分类状态如表 2 所示。数据每段样本的采样点数为 1 024，每种载荷样本数按 3∶1 的比例划分为训练集和测试集。在不同负荷的测试集中加入不同的信噪比成为变噪声的测试样本，实验数据的具体参数如表 3 所示。

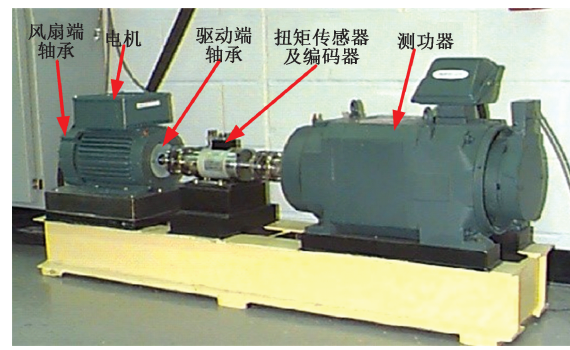


图 5　实验装置图

Fig. 5　Experimental setup diagram

表2　实验故障状态分类

Table 2　Experimental fault state classification

损伤直径/mm	损伤位置	故障状态
0.177	滚动体	故障1
	内圈	故障2
	外圈	故障3
0.355	滚动体	故障4
	内圈	故障5
	外圈	故障6
0.533	滚动体	故障7
	内圈	故障8
	外圈	故障9
0.711	滚动体	故障10
	内圈	故障11
—	—	正常

表3　实验数据的具体参数

Table 3　Specific parameters of experimental data

不同载荷	训练样本数	变噪声测试样本数	变负荷测试样本数		
Ahp	1 149	390(Ahp)	448(Bhp)	448(Chp)	449(Dhp)
Bhp	1 327	448(Bhp)	390(Ahp)	448(Chp)	449(Dhp)
Chp	1 326	448(Chp)	390(Ahp)	448(Bhp)	449(Dhp)
Dhp	1 331	449(Dhp)	390(Ahp)	448(Bhp)	448(Chp)

实验2:采用东南大学变速箱数据集进行实验验证,该实验设备如图6所示,设备主要由电机控制器、电机、减速齿轮箱、行星齿轮箱、负载控制器组成,实验数据包括轴承数据集和齿轮数据集两个子数据集,主要分别在转速-负载为20 Hz-0 V和30 Hz-2 V两种工况下运行,从传动系统动力学模拟器上获取4种故障和1种正常数据集,实验共10种状态。实验数据采集8个通道得到数据,但只采用通道2的数据验证本算法得优越性。此外,数据共采集4 000样本,每个状态由400个样本,按3∶1的比例划分为训练样本和测试样本进行网络训练。

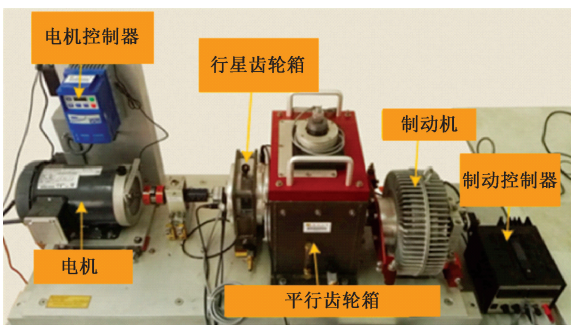


图6　东南大学实验设备

Fig. 6　Diagram of experimental equipment at Southeast University

3.2　实验参数设置

本文改进卷积Lenet-5神经网络具体参数设置如表4所示,表4中输入大小为32×32×1表示将输入一维信号转化成32×32的二维灰度图,通道数设置为1;Block模块中因采用不同尺度的小卷积,参数较多,具体参数如表5所示;输出大小16×16×32表示输出尺寸为16×16,通道数为32;池化1的参数2×2×1表示:池化窗口大小为2×2,步长为1;1 024表示为全连接层输出的神经元的个数,最终输出12种状态的故障。

表4　改进Lenet-5网络具体参数

Table 4　The specific parameters of the improved Lenet-5 network

结构名称	网络层名称	参数	输出大小
输入	—	—	32×32×1
特征提取	Block1模块	—	16×16×32
	池化1	2×2×1	15×15×32
	Block2模块	—	8×8×48
	池化2	2×2×2	4×4×48
	Block3模块	—	2×2×256
全连接层	全连接层	1024	—
分类	—	—	12

Block模块多尺度小卷积具体参数如表5所示,以Block1模块为例:3×3×1×8参数表示卷积核大小为3×3,步长为1,通道数为8;Block2模块卷积核2×2×2的参数表示池化窗口为2×2,步长为2。

表5　Block模块具体参数

Table 5　Specific parameters of Block module

网络层名称	第1个小卷积	第2个小卷积	第3个小卷积
Block1模块	5×5×1×8	1×1×2×8	—
	3×3×1×8	1×1×2×8	—
	1×1×2×8	3×3×1×8	—
	1×1×2×8	—	—
Block2模块	3×3×2×16	—	—
	2×2×2	1×1×1×16	—
	1×1×2×16	—	—
Block3模块	3×3×2×64	1×1×1×64	—
	3×3×1×64	3×3×1×64	1×1×2×64
	1×1×2×64	3×3×1×64	—
	1×1×2×64	—	—

4　实验仿真结果分析

4.1　改进卷积层对模型性能的影响

在实验中,为了验证改进卷积层Block模块对滚动轴承故障分类的结果的影响,引入流行算法t-SNE[25]维数约减算法对3层未改进的卷积层和改进的卷积层特征提取进行可视化,实验结果如图7所示。图7(a)、(b)、(c)分别为未改进Lenet-5三层普通卷积层的可视化图,图7(d)、(e)、(f)分别是改进Lenet-5后Block1模块、Block2模块、Block3模块特征提取可视化图。从图7(a)

和(d)可以看到,经过第 1 层特征提取后输出样本分布情况,图 7(d)Block1 模块所提取的样本大部分都开始趋于各自的区域,而图 7(a)大部分样本还处于分散状态;第 2 层卷积特征提取可视化图 7(b)和(e)可以看到,图 7(e)比 7(b)分类的效果更好,大部分都聚集到各自区域,只有少部分混淆的状态,与第 1 层的特征提取相比较,第 2 层特征提取后的样本聚拢的效果更加明显;经过第 3 层特征提取后,改进 Lenet-5 网络提取的特征分类均聚拢在各自区域上,说明经过 3 层改进的卷积网络所提取的特征信息,已经有较高的故障分辨率。主要是因为改进 Lenet-5 网络卷积层能够自适应标定关键特征信息,减少网络的冗余信息,提高网络特征学习能力,使得高效诊断出故障类型。

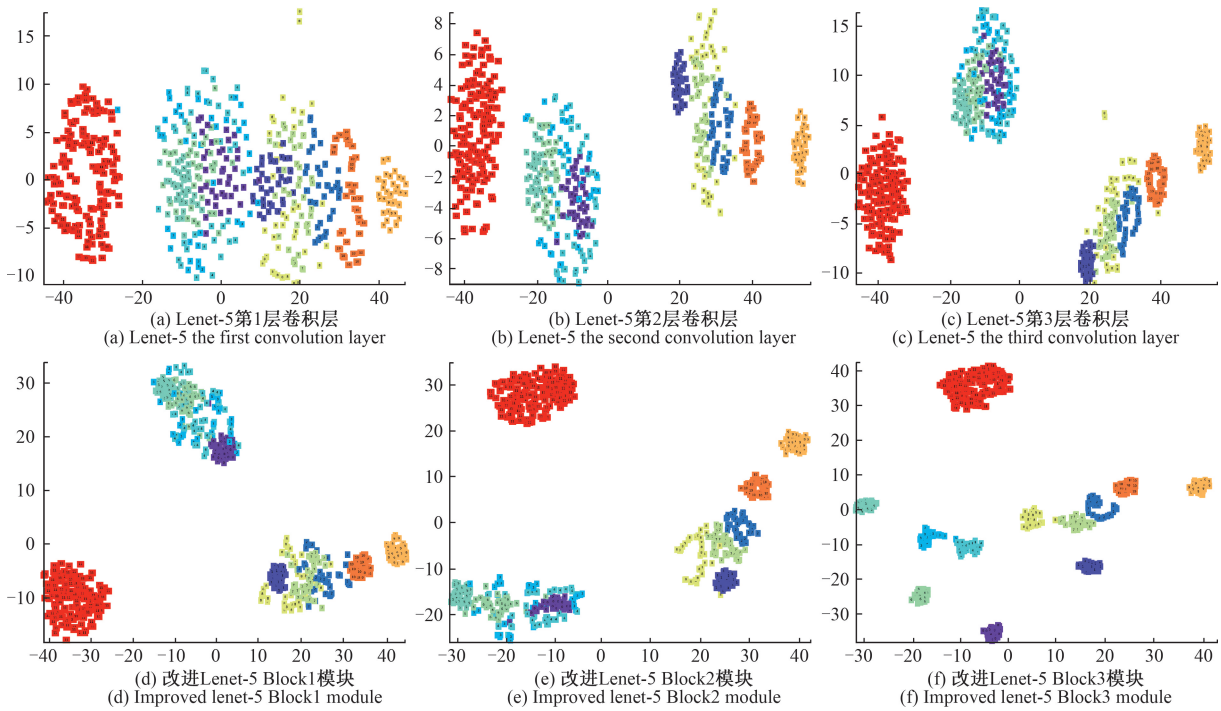


图 7　不同层级可视化图

Fig. 7　Visualization of different levels

4.2　优化网络对模型的影响

为了验证优化网络 L2 正则化和 Dropout 对本文方法优化对比实验效果,采用传统 Lenet-5 与改进的 Lenet-5 方法分别使用优化网络进行网络优化实验对比,实验结果如图 8 所示。横轴表示网络训练迭代的次数,网络每训练一定次数时,随之参数就更新一次,网络训练曲线越稳定表示防止过拟合的效果越好。

从图 8(a)可以看到,训练次数为 300 之后,准确率的曲线基本达到 1,但从曲线可以看到改进卷积网络+L2+Dropout 的准确率曲线上升很平滑,当准确率达到近 100%以后曲线趋于稳定,相对其他曲线很稳定,主要采用 L2 正则化对网络权重参数减小,从而降低网络的复杂度,再由 Dropout 网络对关键信息的神经元激活,从而降低过拟合现象;而改进卷积网络+Dropout 稳定性相对较差,主要是因为 Dropout 网络优化的过程中一些神经元不被激活,使得轴承微弱优质的振动信号特征提取的过程中丢失,因此网络训练所提取的特征信息不能表征轴承实际状态;改进卷积网络+L2 相对改进卷积网络和传统卷积网络的优化后的网络更加稳定。

图 8(b)损失函数曲线图可以看到,当模型训练次数叠加到 500 后,除传统卷积网络,其他网络的损失函数曲线都逐渐趋向于 0;传统卷积网络稳定性差,主要是因为传统卷积网络结构简单,训练过程容易造成过拟合现象,当在网络中采用 L2 正则化和 Dropout 对网络优化后,曲线平滑降低并相对稳定降到趋向于 0,主要是因为优化后的网络所提取的特征信息更加完整、有效,能够表征实际工况下的轴承状态。

4.3　实验 1:滚动轴承故障诊断性能实验对比

为进一步验证本文所构建的模型对故障诊断的性能,选择 Lenet-5、Alexnet、Densenet、改进残差网络[26]作为对比实验进行对比,为了避免实验偶然性,对实验中的每种算法进行 20 次的实验结果求取平均值。其他 3 种比较方法的网络结构具体参数如下:

1) Lenet-5 网络:网络结构与参数已在基础理论阐

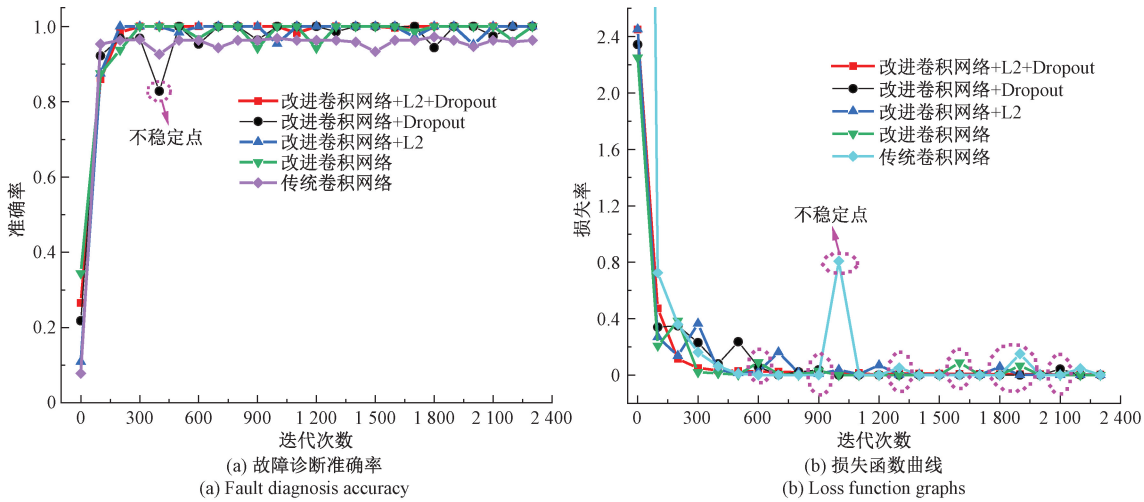


图 8　L2 正则化对网络优化的效果图

Fig. 8　The effect of L2 regularization on network optimization

述,其中网络批尺寸为 64,学习率设置为 0.001,弃权值为 0.5;

2) Alexnet 网络:网络结构为 5 层传统卷积层和 3 层池化层,网络输入的尺寸为 28×28,卷积层的卷积核大小为 3×3,各卷积层的步长分别为 2、2、2、1、2,通道数为 8、16、32、64、128,池化层的通道数和步长都为 1,全连接层数为 1 024;

3) Densenet 网络:网络结构为五层传统卷积层构成的密集网络,网络输入的尺寸为 28×28,卷积层的卷积核大小为 3×3,各卷积的步长都为 2,通道数为 8、8、16、32、160,全连接层数为 1024。

4) 改进残差网路:网络结构由改进的 Inception 模块和带跳跃连接的残差块与空洞残差块构成的改进残差模块组成,Inception 模块采用 3 个 3×3 小卷积核串并联拼成,改进残差块跳跃连接线的参数 λ 取 0.2。

1) 变噪声实验

滚动轴承的振动信号是在机械运转中采集,所采集的信号混杂周边环境的信号,因而具有“强噪声”背景信号,影响轴承故障诊断的识别率。所需的关键信号常常被淹没在噪声种。因此,在信号中加入不同信噪比的高斯白噪声模拟实际中的噪声,验证网络模型在不同噪声下的适应性。选负荷 C 的数据集作为测试集和训练集,并在测试集中加入 6、9、12 dB 不同信噪比的高斯白噪声,进而分析本算法的抗噪性能,与 4 种方法进行比较,实验结果如表 6 所示。

表 6　不同方法变噪声故障诊断率

Table 6　Different methods of variable noise fault diagnosis rate

信噪比	Lenet-5 方法	Alexnet 方法	Densenet 方法	改进残差网络方法	本文方法
6 dB	90.02%	89.55%	90.24%	95.6%	**98.66%**
9 dB	92.46%	95.54%	95.38%	97.58%	**99.1%**
12 dB	95.8%	96.57%	97.26%	98.9%	**99.33%**
平均值	92.76%	93.89%	94.3%	97.36%	**99.03%**

由表 6 故障诊断准确率的对比实验数据可以看到,本文所提的方法不仅在不同信噪比的故障诊断的准确率均高于其他的方法,而且 3 种信噪比准确率的平均值高达 99%以上。基本的 Lenet-5 方法与 Alexnet 方法网络框架相对简单,普通卷积连续单向提取的特征信息学习能力相对差,当信噪比越小,在测试集中的噪声越强时,故障诊断的准确率越低,表明网络抗干扰能力就越差,因此对含有噪声的数据处理能力较弱;而 Densenet 方法故障诊断率相对前面两个基本方法高一些,是因为密集连接网络反复学习前面的特征信息,提取的特征信息比单向网络所提取的特征信息更加充分一些,但由于特征反复利用增加网络训练参数,使网络不能迅速达到最优解,故故障诊断率的平均值才达到 94.3%;改进残差网络采用 Inception 模块与两种改进残差深度特征学习,尽管故障诊断的准确率达 97%,但相对本文方法抗干扰能力相对差些,主要是因为本文采用改进的卷积层和优化网络,可以自适应提取出具有标定的重要特征信息,提高了网络特征学习能力,使得具有更好的抗干扰能力。

2) 变负荷实验

旋转机械的滚动轴承在实际工作中,不仅环境伴有强噪声,而且在不同转速变负荷下运行,因此故障诊断方法是否具有泛化性能对故障诊断具有很重要的作用,为了验证本文所提方法的泛化性能,利用 4 种负荷 Ahp、Bhp、Chp、Dhp,以其中一种负荷为数据样本的训练集,其他负荷的数据样本为测试集进行训练,采用不同方法在变负荷工况下实验对比,实验结果如图 9 所示,图中的 A-B、A-C、A-D 是以 A 负荷数据集为训练集,B、C、D 三种负荷的数据集为测试集进行网络训练得到故障诊断准确率,其他变负荷类似。

由图 9 可以看到本文所提方法在每种负荷的变负荷下,轴承故障诊断率都高于 90%,并且图中的柱高均高于其他方法。传统 Lenet-5 方法采用传统的卷积层所提取的特征学习能力弱,故障诊断的准确率低,基本都在 80% 左右,很少达到 90%,从而对故障诊断的抗泛化性能差;Alexnet 方法与 Densenet 方法相对 lenet-5 方法提取的特征信息更充分,但故障诊断的准确率最高才只有 94.52%,而本文提的方法在变负荷下故障诊断率都高于 90%,而且很多故障诊断率都高于 94% 以上,这是因为本文的方法不仅采用多尺度提取完整的特征信息,更重要的是通过残差收缩网络软阈值化标定重要的特征信息,同时将冗余特征置为 0,得到的特征信息更加利于故障识别,然后利用残差连接加快网络收敛,最终通过 L2 正则化和 Dropout 优化网络,使得网络训练参数不仅减少,还快速找到最优值,自适应不同负荷下的轴承故障诊断,进而提高了网络的泛化能力。

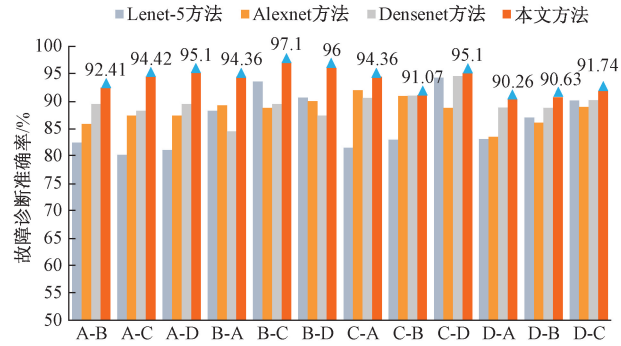


图 9　变负荷工况下实验结果

Fig. 9　Experimental results under variable load conditions

3) 变工况实验

经过前面两个实验验证本文所提方法在单独变噪声和变负荷条件下,具有更好的抗噪性和泛化能力。但实际工业环境,旋转机械工作环境比较复杂,一般在不同载荷运载的同时也伴有强噪声,因此,验证故障诊断的方法同时具备抗噪性与泛化能力,对轴承故障诊断极其必要。本实验采用 4 种不同负荷,其中一种为训练集,其他负荷的数据集为测试集,并且在测试集中加入不同的信噪比(6、9、12 dB)训练,实验的结果如图 10 所示。

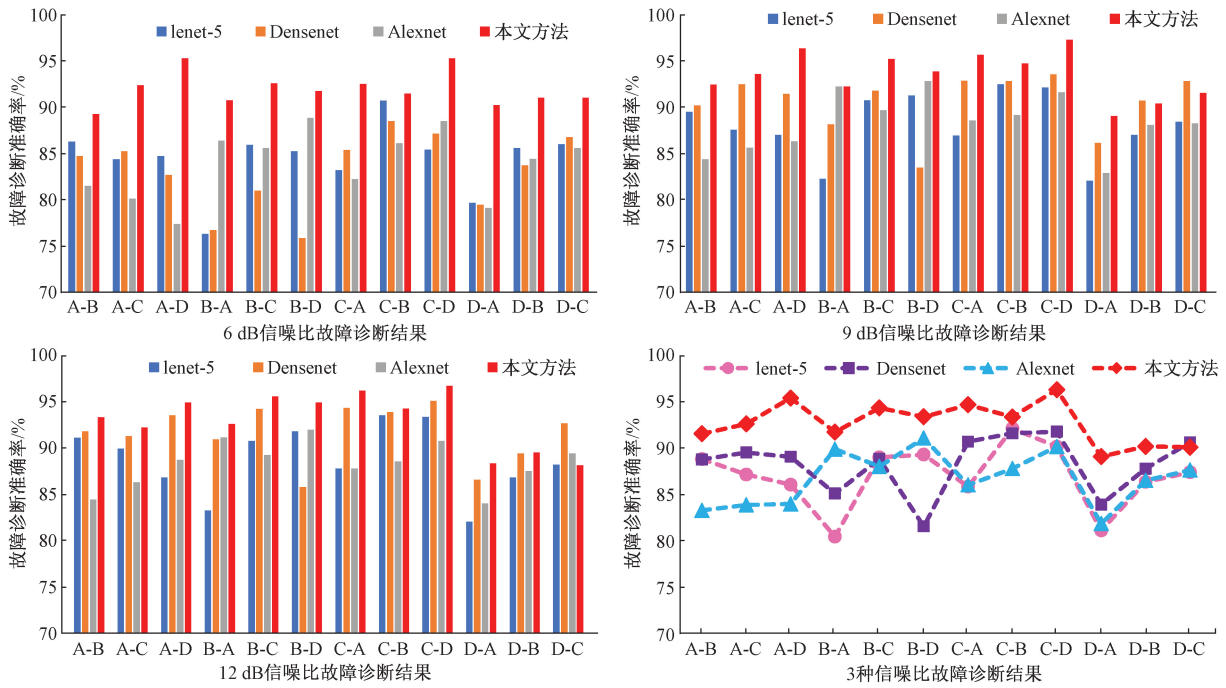


图 10　不同方法变工况故障诊断结果

Fig. 10　Diagram of fault diagnosis results of different methods under variable working conditions

从结果图 10 可以看出,在变噪声与变负荷同时发生的情况下,本文所提的方法对轴承故障诊断的准确率均高于其他 3 种方法,平均值基本都达到 90%以上,并且本文所提的方法在不同信噪比的准确率落差相对较小,平均曲线相对平滑,实验充分说明,改进的卷积层比普通卷积层提取的特征信息能力更高效,更具有表征复杂工况环境的信息。

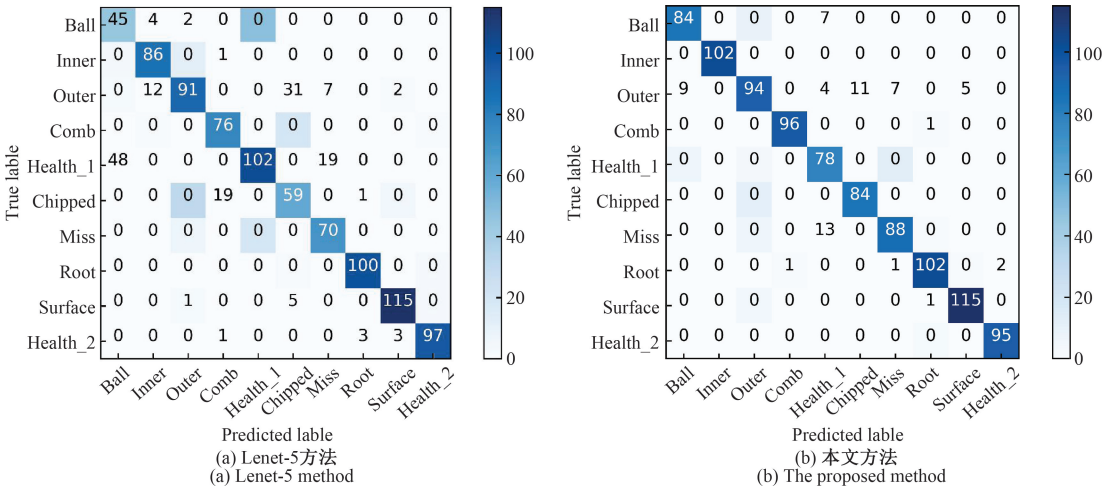


图 11　故障诊断混淆矩阵

Fig. 11　Fault diagnosis confusion matrix diagram

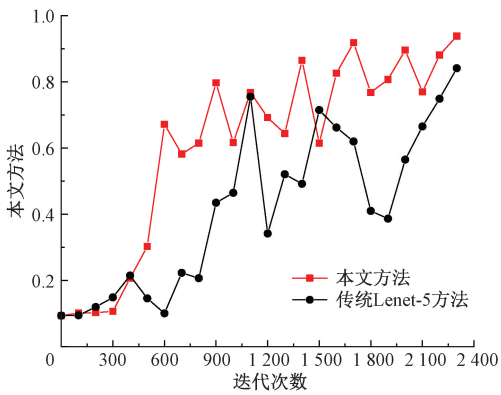


图 12　故障诊断训练结果

Fig. 12　Training results of fault diagnosis

图 11 利用混淆矩阵得到传统 Lenet-5 方法与本文方法的分类结果,结果分别如图 11(a)、(b)所示。图 11 中横坐标和纵坐标分别表示 10 个预测标号和 10 个真实标号,对角线表示每个故障类的正确识别数。可以看到传统 Lenet-5 方法相比本文所提方法的错误分类比较多,这是由于齿轮箱的某些故障特征相似所致,导致它们在识别混合故障时表现不太令人满意。而本文所提的方法误诊断低,说明在识别混合故障方面具有明显的性能优势,能够有效地对混合故障进行诊断。

此外,为了更清晰地验证方法的分类有效性,引入了一些可度量的评价标准。如召回率、精确率、F1 – score,其计算公式如下:

$$召回率 = \frac{TP}{TP + FP} \times 100\% \tag{6}$$

$$精确率 = \frac{TP}{TP + FN} \times 100\% \tag{7}$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \times 100\% \tag{8}$$

其中,TP 表示真阳性样本数,FP 表示假阳性样本数,TN 表示真阴性样本数,FN 表示假阴性样本数。将每一类的结果计算求取平均值如表 7 所示,在表 7 中,本文所提的方法的评价结果均在 93%以上,比传统 Lenet-5 方法高约 10%,并且本文所提方法的所有评价结果都高于传统 Lenet-5 方法,说明利用改进的卷积层优秀的特征学习能力,相比传统卷积层所提取的特征更加的有效,可以提高卷积网络特征提取的效果。

表 7　东南大学数据集诊断结果

Table 7　Diagnostic results of dataset of Southeast University

方法	召回率	精确率	F1-score
传统 Lenet-5	0. 872 4	0. 834 8	0. 835
本文方法	0. 947 9	0. 938	0. 938 6

从图 12 中可以看出,随着迭代次数不断增加训练,本文所提方法不仅准确率高,而且诊断的诊断率更加稳定,充分说明本文所提方法相对于传统的方法所提取的特征信息更加充分,更具有代表齿轮箱运行状态。实验

4.4　实验 2:齿轮箱故障诊断性能实验对比

1)齿轮箱故障诊断

为了验证本文所提方法的有效性,采用东南大学实验数据集对齿轮箱故障诊断进行实验验证,将传统卷积 Lenet-5 方法与本文方法进行对比,诊断的结果如图 11、12 所示。

结果再次说明改进的卷积层特征提取能力比传统卷积层提取的特征更加充分、有效,使得诊断效果更加精确。

2)抗噪性实验结果

在本实验中,为了测试所提方法在混合故障条件下的抗噪声性能,构造了 3 种噪声环境。构建实验环境的步骤与轴承实验相同。采用高斯白噪声模拟环境噪声,只在测试样本中加入噪声信号。实验结果如表 8 所示。可以看出,在 3 种噪声环境下,本文所提的方法的精度高于传统 Lenet-5 方法,这说明本文所提的方法采用改进卷积层在抗噪声性能上具有优势,且具有较高的稳定性。

表 8　东南大学抗噪性实验对比

Table 8　Experimental comparison of anti-noise performance in Southeast University

方法	6 dB	9 dB	12 dB	平均值
Lenet-5 算法	87.69%	90.51%	94.97%	91.06%
本文算法	**92.38%**	**97.83%**	**98.92%**	**96.38%**

5　结　论

本文基于传统卷积提取的特征信息不充分问题,对传统的卷积层进行改进,并应用于简单的 Lenet-5 网络中,实现对强噪声、变负荷下端对端高效的轴承故障诊断,本文得出的结论如下:

1) 改进卷积神经网络:本文结合多尺度卷积方式,将不同尺度提取的特征信息融合,提取更充分、更完整的特征信息;再采用残差收缩网络对冗余信息剔除,保留有益故障分类的特征信息,减少参数的同时,防止过拟合;最后利用残差连接,避免网络训练出现梯度爆炸,加快网络收敛能力,这样组成的 Block 模块,相对传统卷积,提取的特征信息更精细。

2) 随着在网络中将 3 层普通卷积改进为 Block1 模块、Block2 模块、Block3 模块后,网络复杂度提高,导致网络训练出现过拟合现象,因此采用 L2 正则化和 Dropout 进行网络优化,提高网络的特征表达能力。

3) 采用两种不同数据集和不同方法进行比较验证,实验结果表明,本文方法所提的方法对轴承故障诊断更具有抗干扰性和泛化能力。

参考文献

[1] 刘湘楠, 赵学智, 上官文斌. 强背景噪声振动信号中滚动轴承故障冲击特征提取[J]. 振动工程学报, 2021, 34(1): 202-210.
LIU X N, ZHAO X ZH, SHANGGUAN W B. Extraction of impact characteristics of rolling bearing faults in vibration signals with strong background noise [J]. Journal of Vibration Engineering, 2021, 34(1): 202-210.

[2] TRA V, KIM J, KHAN S A, et al. Bearing fault diagnosis under variable speed using convolutional neural networks and the stochastic diagonal Levenberg marquardt algorithm[J]. Sensors, 2017, 17(12): 2834.

[3] CIABATTONI L, FERRACUTI F, FREDDI A, et al. Statistical spectral analysis for fault diagnosis of rotating machines [J]. IEEE Transactions on Industrial Electronics, 2018, 65(5): 4301-4310.

[4] VIET T, JAEYOUNG K, ALI K S, et al. Bearing fault diagnosis under variable speed using convolutional neural networks and the stochastic diagonal levenberg-marquardt algorithm [J]. Sensors (Basel), 2017, 7(12): 2834-2849.

[5] 金国强. 基于深度学习的复杂工况下端到端的滚动轴承故障诊断算法研究[D]. 合肥:中国科学技术大学, 2020.
JIN G Q. Research on end-to-end rolling bearing fault diagnosis algorithm under complex working conditions based on deep learning[D]. Hefei: University of Science and Technology of China, 2020.

[6] 丁承君, 冯玉伯, 王曼娜. 基于变分模态分解与深度卷积神经网络的滚动轴承故障诊断[J]. 振动与冲击, 2021, 40(2): 287-296.
DING CH J, FENG Y B, WANG M N. Rolling bearing fault diagnosis based on variational mode decomposition and deep convolutional neural network[J]. Vibration and Shock, 2021, 40(2): 287-296.

[7] 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.

[8] 叶壮, 余建波. 基于多通道加权卷积神经网络的齿轮箱振动信号特征提取[J]. 机械工程学报, 2021, 57(1): 110-120.
YE ZH, YU J B. Gearbox vibration signal feature extraction based on multi-channel weighted convolutional neural network [J]. Chinese Journal of Mechanical Engineering, 2021, 57(1): 110-120.

[9] 武哲, 杨绍普, 刘永强. 基于多元经验模态分解的旋转机械早期故障诊断方法[J]. 仪器仪表学报, 2016, 37(2): 241-248.
WU ZH, YANG SH P, LIU Y Q. Early fault diagnosis method for rotating machinery based on multiple empirical mode decomposition [J]. Chinese Journal of Scientific Instrument, 2016, 37(2): 241-248.

[10] 董绍江, 裴雪武, 吴文亮, 等. 基于多层降噪技术及改进卷积神经网络的滚动轴承故障诊断方法[J]. 机械工程学报, 2021, 57(1): 148-156.
DONG SH J, PEI X W, WU W L, et al. Rolling bearing fault diagnosis method based on multi-layer noise reduction technology and improved convolutional neural network[J]. Chinese Journal of Mechanical Engineering, 2021, 57(1): 148-156.

[11] YE T, JIAN M, CHEN L, et al. Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine[J]. Mechanism and Machine Theory, 2015, 90(8): 175-186.

[12] 徐彦伟, 蔡薇薇, 颉潭成, 等. 变工况下基于信息融合的地铁牵引电机轴承故障智能诊断[J]. 计算机集成制造系, 2021, 36(4):1-20.
XU Y W, CAI W W, JIE T CH, et al. Intelligent diagnosis of subway traction motor bearing faults based on information fusion under variable working conditions[J]. Computer Integrated Manufacturing System, 2021, 36(4): 1-20.

[13] 杨洋, 黄罗杰, 李平, 等. 基于多维度特征提取的电弧故障检测方法[J]. 电子测量与仪器学报, 2021, 35(10):107-115.
YANG Y, HUANG L J, LI P, et al. A multi-dimensional feature extraction-based method for arc fault detection[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(10):107-115.

[14] ZHANG Z Y, WANG Y, WANG K S. Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network[J]. Journal of Intelligent Manufacturing, 2013, 24(6): 1213-1227.

[15] 许子非, 金江涛, 李春. 基于多尺度卷积神经网络的滚动轴承故障诊断方法[J]. 振动与冲击, 2021, 40(18):212-220.
XU Z F, JIN J T, LI CH. Rolling bearing fault diagnosis method based on multi-scale convolutional neural network[J]. Journal of Vibration and Shock, 2021, 40(18): 212-220.

[16] PENG D, WANG H, LIU Z, et al. Multibranch and multiscale CNN for fault diagnosis of wheelset bearings under strong noise and variable load condition[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4949-4960.

[17] 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, 2017, 99:1.

[18] 车畅畅, 王华伟, 倪晓梅, 等. 基于深度残差收缩网络的滚动轴承故障诊断[J]. 北京航空航天大学学报, 2020,46(11): 1-10.
CHE CH CH, WANG H W, NI X M, et al. Rolling bearing fault diagnosis based on deep residual shrinkage network[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(11): 1-10.

[19] 赵小强, 张青青. 改进 Alexnet 的滚动轴承变工况故障诊断方法[J]. 振动. 测试与诊断, 2020, 40(3): 472-480,623.
ZHAO X Q, ZHANG Q Q. Improved Alexnet's fault diagnosis method for rolling bearing under variable conditions[J]. Vibration, Test and Diagnosis, 2020, 40(3): 472-480,623.

[20] 郭晨, 简涛, 徐从安, 等. 基于深度多尺度一维卷积神经网络的雷达舰船目标识别[J]. 电子与信息学报, 2019, 41(6): 1302-1309.
GUO CH, JIAN T, XU C AN, et al. Radar ship target recognition based on deep multi-scale one-dimensional convolutional neural network[J]. Journal of Electronics & Information Technology, 2019, 41(6): 1302-1309.

[21] 仝钰, 庞新宇, 魏子涵. 基于 GADF-CNN 的滚动轴承故障诊断方法[J]. 振动与冲击, 2021, 40(5): 247-253,260.
TONG Y, PANG X Y, WEI Z H. Fault diagnosis method of rolling bearing based on GADF-CNN[J]. Vibration and Shock, 2021, 40(5): 247-253,260.

[22] 樊湘鹏, 周建平, 许燕, 等. 基于改进卷积神经网络的复杂背景下玉米病害识别[J]. 农业机械学报, 2021, 52(3): 210-217.
FAN X P, ZHOU J P, XU Y, et al. Recognition of corn diseases in complex background based on improved convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Machinery, 2021, 52(3): 210-217.

[23] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212-223.

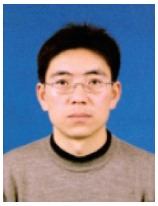
[24] SMITH W A, RANDALL R B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study[J]. Mechanical Systems and Signal Processing, 2015, 64: 100-131.

[25] SPIWOK V, KRÍŽ P. Time-lagged t-distributed stochastic neighbor embedding (t-SNE) of molecular simulation trajectories. [J]. Frontiers in Molecular Biosciences, 2020:132.

[26] 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 X Q, LIANG H P. Fault diagnosis method of

rolling bearing under variable conditions using improved residual neural network[J]. Journal of Xi′an Jiaotong University, 2020, 54(9): 23-31.

作者简介



赵小强,2006 年于浙江大学获得博士学位,现为兰州理工大学教授、博士生导师,主要研究领域为故障诊断、图像处理、数据挖掘等。

E-mail: xqzhao@lut.cn

Zhao Xiaoqiang received the Ph. D. degree from Zhejiang University in 2006. Now he is a professor and Ph. D. supervisor at Lanzhou University of Technology. His main research interests include fault diagnosis, image processing and data mining etc.

罗维兰,硕士研究生,主要研究方向为故障诊断。

E-mail: 2700268805@qq.com

Luo Weilan is a M. Sc. candidate. Her main research interest includes fault diagnosis.