

# 面向数据不平衡的卷积神经网络故障辨识方法\*

吴耀春<sup>1,2</sup>, 赵荣珍<sup>1</sup>, 靳伍银<sup>1</sup>, 邢自扬<sup>1</sup>

(1. 兰州理工大学机电工程学院 兰州, 730050) (2. 安阳工学院机械工程学院 安阳, 455000)

**摘要** 针对因不同故障的样本数目不平衡造成卷积神经网络(convolutional neural network, 简称 CNN)对少数类样本识别准确率偏低的缺陷, 采用将一种最小最大化目标函数融入卷积神经网络结构的对策, 提出一种适用于故障数据不平衡的最小最大化目标函数卷积神经网络(min-max objective CNN, 简称 MMOCNN)智能故障模式辨识方法。首先, 利用卷积神经网络交替的卷积与池化运算自适应学习振动信号中具有表征信息的敏感特征, 并通过全连接层(fully connected layer, 简称 FC)将学习特征映射到类空间; 其次, 在类空间构造特征的最小最大化目标函数; 最后, 将最小最大化目标函数融入到卷积神经网络的损失函数中, 在模型训练过程中既考虑分类总体误差最小, 同时又要求学习的样本特征保持同类距离小、异类距离大, 以实现数据不平衡故障的有效辨识。用轴承的不平衡数据集分别对本方法和传统卷积神经网络的辨识效果进行实验, 结果表明, 本方法能够使少数类样本的辨识精度提升 20% 以上。

**关键词** 智能故障辨识; 数据不平衡; 卷积神经网络; 最小最大化目标函数; 旋转机械

**中图分类号** TH165.3; TP206.3

## 引言

旋转机械在现代工业生产中的作用举足轻重<sup>[1-2]</sup>。随着传感器、计算机与通信、数据存储等技术的高速发展, 旋转机械向大功率复杂化方向发展, 为保障此类设备安全、优质高效运行, 其监测装置的规模正在逐渐扩大, 导致旋转机械监测装置获取的数据呈现出海量、复杂、高维、多源异构和价值低密度等大数据特性<sup>[3-4]</sup>。因此, 改进传统基于信号处理的特征提取+机器学习的方法性能, 发展基于工业大数据资源开发利用的旋转机械智能决策技术, 对于加快发展智能旋转机械将具有非常重要的推动作用<sup>[5-6]</sup>。

近年来, 作为大数据挖掘利器的深度学习, 因其具有可利用原始观测数据、通过自动学习方式即可得到故障的更有效表征信息和能够提取出敏感特征的能力, 典型的如 CNN 网络等, 已经成为旋转机械智能故障辨识研究的热点<sup>[7-9]</sup>。Janssens 等<sup>[10]</sup>建立了由卷积层和 FC 构成的浅层卷积神经网络模型, 以轴承原始振动信号的频谱作为数据源, 获得了 93.61% 的故障识别准确率。Jing 等<sup>[11]</sup>建立了一种基于 CNN 的特征学习与故障诊断方法, 该方法在齿轮箱故障诊断的测试中达到 98% 的辨识精度。Wen 等<sup>[12]</sup>将振动信号转换为 2D 图像, 在 LeNet-5 的

基础上构造出一种改进 CNN 模型用于轴承故障的诊断。Zhang 等<sup>[13]</sup>利用一种新型训练方法训练深度 CNN 模型, 实现了强背景噪声、变载荷条件下的轴承故障有效识别。

基于 CNN 模型原理的深度学习方式的性能仍需进一步完善, 其缺陷之一是因忽视样本分布不平衡对故障辨识准确率的影响作用。机械故障是偶发的, 故障状态信号相较于正常运行数据更加难以采集, 因此获取的故障数据具有明显的不平衡性。基于 CNN 的智能故障诊断方法通常是以提高分类总体精度为目标, 数据的不平衡会迫使这些方法偏向于多数类样本的特征学习, 而对少数类样本的学习不足, 容易造成分类器对少数类样本识别精度下降<sup>[14]</sup>。然而在工业设备监测中, 机械设备故障作为少数类发生, 要比多数类(正常)更加重要, 若将“故障”误诊为“正常”, 让机械系统继续工作, 会导致无法预计的后果和损失<sup>[15]</sup>。因此, 在数据不平衡条件下如何准确辨识判断机械系统运行状态是关键问题。

目前, 解决数据不平衡分类问题的方法可以分为数据层面和算法层面两大类。数据层面方法主要有欠采样<sup>[16]</sup>、过采样<sup>[17]</sup>和合成少数类过采样(synthetic minority oversampling technique, 简称 SMOTE)<sup>[18]</sup>。欠采样是通过减少多数类样本的数量, 过采样是增加少数类样本的数量, 从而使得样本的分布达到平衡。

\* 国家自然科学基金面上资助项目(51675253); 国家重点研发计划资助项目(2016YFF0203303-04); 河南省重点研发与推广专项资助项目(222102220092)

收稿日期: 2020-08-18; 修回日期: 2020-10-25

本质上这两种方法并没有增加额外的数据信息,并且欠采样通过样本剔除有可能破坏数据集的分布,而过采样有可能引入噪声造成模型的过拟合。SMOTE采样通过随机选择同类近邻样本进行插值,生成无重复新的少数类样本,并出现了一系列SMOTE改进算法<sup>[19]</sup>,但由于合成噪声和边界样本,它们的扩展决策边界仍然难以区分。算法层面方法主要有集成学习<sup>[20]</sup>和代价敏感学习<sup>[14,21]</sup>。集成学习是通过将多个基分类器在平衡数据子集上的分类结果按一定方式集成来提升分类器的泛化性能,但这种方法的效率较低。代价敏感学习通过给少数类样本分配更高的错分代价,以分类错误总代价最低为优化目标来处理数据不平衡问题,这种方法简单直观但需要预先定义代价矩阵。事实上,分类器在数据分布边界清晰的不平衡数据集上也可以得到较好的分类结果<sup>[22]</sup>。在面对样本空间分布重叠、特征不明显的不平衡故障数据集,如果通过有目的特征学习能够获得类边界清晰的特征,就可缓解由于机械数据不平衡导致少数类样本分类正确率偏低的问题。

基于上述分析,针对机械系统健康状况监测中呈现的正常数据丰富、故障数据匮乏现象,以及由此导致的卷积神经网络在处理这种不平衡数据分类时对少数类样本识别率较低的问题,对改进卷积神经网络性能使其避免因数据不平衡干扰智能故障辨识准确率的方法进行研究。

## 1 CNN原理与不平衡分类问题简介

### 1.1 CNN原理

CNN是一种深度前馈神经网络,它的基本组成包括卷积层、池化层和FC。其中,卷积层和池化层一般用于提取特征,FC利用提取的特征实现分类。由于本研究的输入数据是一维振动信号,这里仅介绍一维CNN。

卷积层采用多卷积核对输入数据进行卷积运算并输出特征矢量,每一个卷积核输出一个特征矢量,对应提取一种特征。这个特征矢量的数学模型可表示为

$$x_j^l = \text{ReLU} \left( \sum_{i \in M_j} x_i^{l-1} k_{ij}^l + b_j^l \right) \quad (1)$$

其中: $x_j^l$ 为第 $l$ 层的第 $j$ 个卷积核输出的特征矢量; $M_j$ 为第 $l-1$ 层输出的特征矢量集合; $x_i^{l-1}$ 为第 $l-1$ 层输出的第 $i$ 个特征矢量; $k_{ij}^l$ 表示第 $l-1$ 层的第 $i$ 个特征矢量与第 $l$ 层的第 $j$ 个特征矢量之间的卷积核矩阵; $b_j^l$ 为第 $l$ 层的第 $j$ 个特征矢量对应的偏置; $\text{ReLU}(\cdot)$ 为激活函数。

卷积层之后是池化层,对输入的特征矢量进行降采样,在实现数据降维的同时进一步突出提取的

特征。本研究采取的最大池化运算公式为

$$p_j^l = \max \{ x_{j \times S:(j+1) \times S}^{l-1} \} \quad (2)$$

其中: $p_j^l$ 为池化运算中第 $l$ 层的第 $j$ 个神经元输出值; $\max \{ \cdot \}$ 为池化运算方式,即选取计算区域的最大值作为输出; $x_{j \times S:(j+1) \times S}^{l-1}$ 为第 $l-1$ 层的第 $j$ 个局部区域; $S$ 为该区域的宽度,即池化尺寸。

FC是卷积池化部分与分类器的连接过渡结构,它通常由一个2~3层的全连接前馈神经网络组成。多次卷积池化运算后,网络学习的特征矢量集合经过Flatten展平函数转变为一维特征矢量,在经过FC时映射到类空间,其计算公式为

$$x^l = \text{ReLU} \left( \sum w_j^l x^{l-1} + b_j^l \right) \quad (3)$$

其中: $x^{l-1}$ 为第 $l-1$ 层的输出值; $w_j^l$ 为第 $l-1$ 层和第 $l$ 层之间的全连接权值矩阵; $b_j^l$ 为第 $l$ 层的偏置向量。

### 1.2 不平衡分类问题

CNN不能有效解决数据不平衡分类问题的原因在于其损失函数是以分类的总体精度最高为优化目标。CNN使用的softmax交叉熵损失函数为

$$L_s = - \sum_{i=1}^C 1 \{ y_i = c \} \log(p_i) \quad (4)$$

$$p_i = e^{x_i} / \sum_{k=1}^C e^{x_k} \quad (5)$$

其中: $x_i$ 为FC的第 $i$ 个神经元输出值; $p_i$ 为FC第 $i$ 个神经元输出的概率值; $C$ 为类别数; $c$ 为输入样本真实健康状态; $y_i$ 为网络对输入样本预测值; $1 \{ \cdot \}$ 表示大括号内等式成立时为1,否则为0。

在模型训练过程中,交叉熵损失函数主要考虑每一个样本对应于其本应所属类别的概率。如果该概率较低,说明模型对该样本特征学习程度不够,予较大的损失;反之,则赋予该样本较小的损失。但对于不平衡数据集而言,由于少数类样本占比小,样本的总体识别率主要由多数类样本的识别率主导。因此,多数类样本误分类对CNN的训练过程影响较大,而少数类样本误分类在一定程度上被忽略,这导致CNN对少数类样本的特征学习不充分,在分类器上显现出少数类样本错分率高。

## 2 数据不平衡故障诊断模型的设计

### 2.1 最小最大化目标函数

最小最大化目标是直接要求CNN在第 $l$ 层提取的样本特征矢量同类距离最小,异类距离最大<sup>[23]</sup>,以利于获取清晰的样本特征分类边界。假设标签为 $Y_Q = \{ y_1, y_2, \dots, y_Q \}$ 的最小批次样本输入卷积神经网络,在第 $l$ 层提取的样本特征集为

$X_Q = \{x_1, x_2, \dots, x_Q\}$ , 其中:  $Q$  为训练样本的最小批次;  $x_i$  为第  $i$  个样本的特征矢量集合展平。最小批次样本特征的类内邻接矩阵的元素  $A_{i,j}^W$  和类间邻接矩阵的元素  $A_{i,j}^B$  可定义<sup>[24]</sup>为

$$A_{i,j}^W = \begin{cases} 1/Q_c & (y_i = y_j = c) \\ 0 & (\text{其他}) \end{cases} \quad (6)$$

$$A_{i,j}^B = \begin{cases} 1/Q - 1/Q_c & (y_i = y_j = c) \\ 1/Q & (\text{其他}) \end{cases} \quad (7)$$

其中:  $Q_c$  为最小批次样本中第  $c$  类样本的个数。

式(6)、式(7)分别给出了最小批次样本特征的类内邻接矩阵元素和类间邻接矩阵元素的定义。为了在计算机平台算法实现的方便性,最小批次样本的类内邻接矩阵  $A^W$  和类间邻接矩阵  $A^B$  可以由式(8)、式(9)来计算

$$A^W = (A_{i,j}^W)_{Q \times Q} = (Y_Q Y_Q^T) [\text{diag}(\epsilon + \sum_{i=1}^Q Y_Q)]^{-1} \quad (8)$$

其中:  $Y_Q$  为最小批次样本标签的独热编码;  $\text{diag}(\cdot)$  为对角矩阵;  $\epsilon$  为  $1 \times C$  的向量, 元素取  $10^{-8}$ 。

$$A^B = (A_{i,j}^B)_{Q \times Q} = \left[ \frac{1}{Q} \right]_{Q \times Q} - (Y_Q Y_Q^T) [\text{diag}(\epsilon + \sum_{i=1}^Q Y_Q)]^{-1} \quad (9)$$

其中:  $\left[ \frac{1}{Q} \right]_{Q \times Q}$  是元素为  $\frac{1}{Q}$  的  $Q \times Q$  矩阵。

最小批次样本特征的类内散度矩阵  $S^W$  和类间散度矩阵  $S^B$  可以由式(10)、式(11)来计算

$$S^W = \frac{1}{2} \sum_{i,j=1}^Q A_{i,j}^W (x_i - x_j)(x_i - x_j)^T \quad (10)$$

$$S^B = \frac{1}{2} \sum_{i,j=1}^Q A_{i,j}^B (x_i - x_j)(x_i - x_j)^T \quad (11)$$

根据式(10)、式(11),最小批次样本特征的同类距离  $S_W$  和异类距离  $S_B$  可以表示为

$$S_W = \text{tr}(S^W) = \frac{1}{2} \sum_{i,j=1}^Q A_{i,j}^W \|x_i - x_j\|^2 = \mathbf{1}_Q^T (A^W \odot \Phi) \mathbf{1}_Q \quad (12)$$

$$S_B = \text{tr}(S^B) = \frac{1}{2} \sum_{i,j=1}^Q A_{i,j}^B \|x_i - x_j\|^2 = \mathbf{1}_Q^T (A^B \odot \Phi) \mathbf{1}_Q \quad (13)$$

其中:  $\text{tr}(\cdot)$  表示矩阵的迹; 矩阵  $\Phi = (\Phi_{i,j})_{Q \times Q}$ ,  $\Phi_{i,j} = \|x_i - x_j\|^2$ ;  $\odot$  表示矩阵点乘运算;  $\mathbf{1}_Q \in R^Q$  表示所有元素为 1 的列向量。

由式(12)、式(13)可以看出,最小化  $S_W$  可以使同类之间的样本特征紧凑,最大化  $S_B$  可以使异类之间的样本特征分离。因此,最小最大化目标函数定义为

$$L_{\text{MMO}} = S_B / S_W \quad (14)$$

由于

$$\frac{\partial S_W}{\partial x_i} = \frac{\partial}{\partial x_i} \left( \frac{1}{2} \sum_{i,j=1}^Q A_{i,j}^W \|x_i - x_j\|^2 \right) = (x_i \mathbf{1}_Q^T - H)(A^W + (A^W)^T)_{(:,i)} \quad (15)$$

$$\frac{\partial S_B}{\partial x_i} = \frac{\partial}{\partial x_i} \left( \frac{1}{2} \sum_{i,j=1}^Q A_{i,j}^B \|x_i - x_j\|^2 \right) = (x_i \mathbf{1}_Q^T - H)(A^B + (A^B)^T)_{(:,i)} \quad (16)$$

其中:  $H = [x_1, x_2, \dots, x_Q]$ , 下标符号  $(:, i)$  表示矩阵的第  $i$  列。

最小最大化目标函数对样本特征向量  $x_i$  的偏导数为

$$\frac{\partial L_{\text{MMO}}}{\partial x_i} = \frac{\frac{\partial S_B}{\partial x_i} S_W - S_B \frac{\partial S_W}{\partial x_i}}{S_W^2} \quad (17)$$

## 2.2 融入最小最大化目标函数的卷积神经网络设计

在面对数据不平衡的故障辨识问题时,传统的 CNN 算法存在着对少数类样本识别率偏低的缺陷。导致这种现象的直接原因是 CNN 采用 softmax 交叉熵作为损失函数进行网络参数优化,模型在训练的过程中偏向多数类样本特征学习,而忽略了少数类样本特征学习。针对此问题,本小节拟将最小最大化目标函数融入到 CNN 的损失函数中,提出 MMOCNN 模型,如图 1 所示。模型重新构造的损失函数为

$$L(\theta) = L_s - \alpha L_{\text{MMO}} \quad (18)$$

其中:  $\theta$  为网络参数集合;  $L_s$  为交叉熵损失函数;  $L_{\text{MMO}}$  为最小最大化目标函数;  $\alpha$  为调整参数。

MMOCNN 的损失函数在考虑分类总体精度的同时,要求模型提取样本特征的  $S_W$  小,  $S_B$  大,期待能够更充分地学习少数类样本特征,以提高模型对不平衡数据中少数类样本的识别精度。

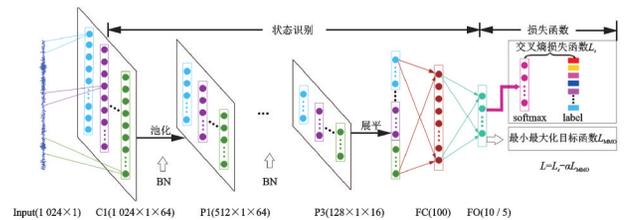


图 1 MMOCNN 结构

Fig.1 The structure of MMOCNN

MMOCNN 模型参数的设计主要是卷积核的个数和尺寸。文献[8]指出,在设置深度学习网络的隐层神经元数目时,需遵循的基本原则应当是后一层神经元数目不超过前一层神经元数目的一半,故本研究将 3 层卷积层的神经元数目分别设为 64, 32 和 16。卷积核的尺寸第 1 层采用大尺寸  $16 \times 1$  的卷积

核,这样可以更好地抑制振动信号中的高频噪声<sup>[13]</sup>。后面的卷积层均采用小尺寸 $3 \times 1$ 的卷积核,便于网络加深,提高网络的学习能力。另外,为了加快收敛速度和提升泛化能力,笔者在模型的卷积层后尝试引入批归一化处理,在FC使用dropout处理。

模型训练主要是通过最小化损失函数来优化网络参数。根据定义的损失函数采用最小批次梯度下降法(stochastic gradient descent,简称SGD)优化网络中的参数,具体的过程为

$$\theta \leftarrow \theta - \eta \frac{\partial L(\theta)}{\partial \theta} \quad (19)$$

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{\partial L_s}{\partial \theta} - \alpha \frac{\partial L_{MMO}}{\partial \theta} \quad (20)$$

其中: $\eta$ 为学习率; $\partial L_s / \partial \theta$ 为交叉熵损失函数的梯度; $\partial L_{MMO} / \partial \theta$ 为最小最大化目标函数的梯度,由式(17)计算。

实验中发现,最小批次 $Q$ 设置过小,图1模型的损失函数振荡严重,在最大迭代次数下难以收敛;批次设置过大,影响模型泛化能力。经过反复调试,最小批次为28,最大迭代次数设为100,学习率 $\eta$ 设为0.001。模型的具体参数如表1所示。

表1 MMOCNN模型的具体参数

Tab.1 Special parameters of MMOCNN

结构	尺寸/步距	输出尺寸	训练参数
C1(卷积)	$16 \times 1 / 1 \times 1$	$1024 \times 1 @ 64$	SGD
P1(最大池化)	$2 \times 1 / 2 \times 1$	$512 \times 1 @ 64$	minibatch=28
C2(卷积)	$3 \times 1 / 1 \times 1$	$512 \times 1 @ 32$	$\eta=0.001$
P2(最大池化)	$2 \times 1 / 2 \times 1$	$256 \times 1 @ 32$	max epoch=100
C3(卷积)	$3 \times 1 / 1 \times 1$	$256 \times 1 @ 16$	dropout=0.3
P3(最大池化)	$2 \times 1 / 2 \times 1$	$128 \times 1 @ 16$	
FC(全连接)	$1 \times 1$	100	
FO(全连接)	$1 \times 1$	C	

### 2.3 归纳的诊断方法与流程

在图1所示的MMOCNN模型中将最小最大化目标函数融入到交叉熵损失函数中,在模型训练时既考虑分类总体精度高,又要求网络学习样本特征的 $S_w$ 小、 $S_B$ 大。将该模型应用于旋转机械数据不平衡故障辨识的具体流程如图2所示。

### 2.4 采用的评价指标

实验采用混淆矩阵和准确率作为模型分类性能的评价指标。混淆矩阵的行为测试样本的模型预测结果,列为测试样本的真实标签,对角线表示模型预测值和样本标签一致的样本数目。定义准确率为

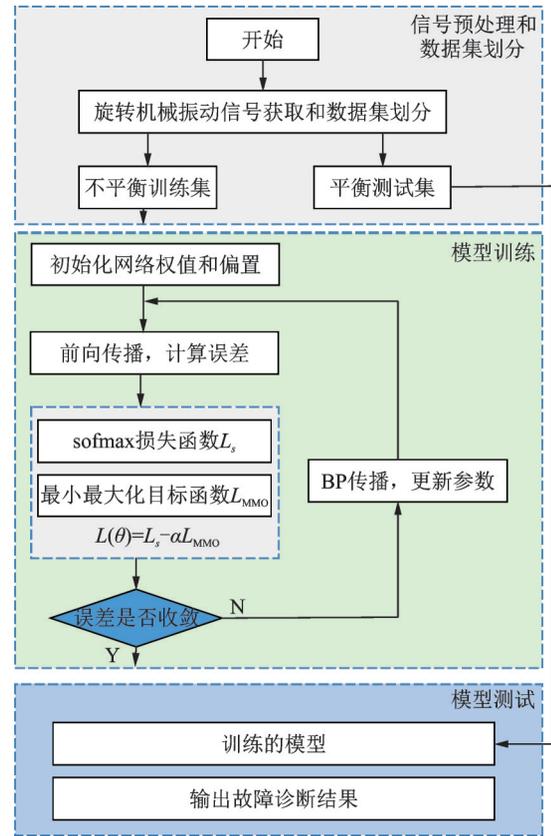


图2 所提方法的流程图

Fig.2 Flow chart of the proposed method

$$\text{Accuracy} = \sum_{i=1}^N \frac{t_i}{p_i} / N \quad (21)$$

其中: $t_i$ 为第 $i$ 次实验诊断正确样本的个数; $p_i$ 为第 $i$ 次实验测试样本的总数; $N$ 为实验次数。

## 3 实验结果与分析

为验证模型(图1)的有效性,本节使用美国凯斯西储大学电气工程实验室的滚动轴承数据进行实验<sup>[25]</sup>。数据采集于机械系统驱动端的加速度传感器,负载为0.746 kW,采样频率为12 kHz。数据集中除了正常状态外,还包括轴承的3种故障状态:滚动体损伤、内圈故障、外圈故障。每一种故障的损伤直径分别为0.177 8 mm(轻微故障)、0.355 6 mm(中等故障)和0.533 4 mm(严重故障)。因此,组合之后的数据集中包括轴承10种不同运行状态,即{BS, IRS, ORS, BM, IRM, ORM, BL, IRL, ORL, NR}。故障类型与符号之间的对应关系为滚动体轻微故障(BS)、内圈轻微故障(IRS)、外圈轻微故障(ORS)、滚动体中等故障(BM)、内圈中等故障(IRM)、外圈中等故障(ORM)、滚动体严重故障(BL)、内圈严重故障(IRL)、外圈严重故障(ORL)和正常状态(NR)。每一种健康状态采集300个样本,样本的长度为1024点。

根据划分的轴承运行状态,建立用于实验的数据集,见表2。 $A$ 为平衡数据集,每一种状态的训练样本均为200个。 $B$ 为不平衡数据集,在设备状态监测的实际应用中,故障状态数据较之正常运行数

据更难以获取,严重故障数据较之轻微故障数据更难以获取,因而轻微、严重故障状态的样本个数依次减少,以模拟实际应用中状态数据的不平衡。

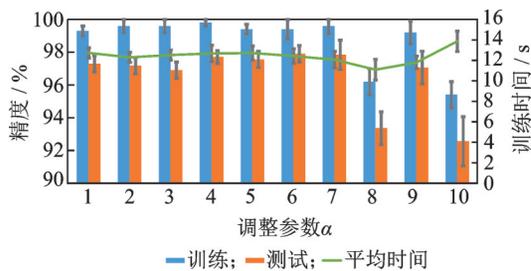
表2 不平衡数据集的描述

Tab.2 Description of imbalanced dataset

数据集		运行状态									
		NR	BS	IRS	ORS	BM	IRM	ORM	BL	IRL	ORL
$A$	训练样本	200	200	200	200	200	200	200	200	200	200
	测试样本	100	100	100	100	100	100	100	100	100	100
$B$	训练样本	200	100	100	100	50	50	50	20	20	20
	测试样本	100	100	100	100	100	100	100	100	100	100

### 3.1 调整参数 $\alpha$ 的选择

在2.2节模型(图1)的优化过程中,两种损失函数的调整参数 $\alpha$ 值的设置对模型最终诊断精度影响较大。笔者采用网格搜索方法确定 $\alpha$ 的值,寻优范围为 $\{1 \times 10^{-6}, 2 \times 10^{-6}, \dots, 10 \times 10^{-6}\}$ ,用数据集 $B$ 测试模型的训练精度、测试精度和训练时间,每个实验重复10次,得到的结果如图3所示。在考虑模型的稳定性和训练难易程度后,最优的 $\alpha$ 取值 $6 \times 10^{-6}$ 。

图3 参数 $\alpha$ 对MMOCNN性能的影响Fig.3 The effect of  $\alpha$  on the performance of MMOCNN

### 3.2 最小最大化目标函数的有效性分析

为验证式(14)所示的最小最大化目标函数的有效性,按照图2所示诊断流程,用表2构造数据集的训练集分别训练MMOCNN和CNN,用测试集分别对训练的模型进行测试,结果如图4所示。

由图4(a,b)可看出:用数据集 $A$ 中的平衡训练集训练模型时,CNN将3个BL类样本错误识别为BS类,7个ORL类样本分别错误识别为BS,ORS和IRS类,总体识别率为99%;MMOCNN在测试集上没有错误识别样本,总体识别率为100%,这说明MMOCNN和CNN在平衡数据上均具有较高的辨识精度。由图4(c,d)可看出:用数据集 $B$ 中的不平衡训练集训练模型时,CNN将31个少数类BL样本错误识别为BS类,54个少数类ORL样本分别错误识别为

IRS,ORS和ORM类,模型的总体识别率为91.5%,但对少数类样本的识别率为71.67%;MMOCNN将13个少数类BL样本错分,8个少数类ORL样本错分,模型的总体识别率为97.7%,对少数类样本的识别率为92.66%,这说明数据不平衡对CNN和MMOCNN模型的诊断性能均有一定的影响,MMOCNN模型在不平衡数据上的诊断性能明显优于CNN,总体诊断精度提高6.2%,少数类样本诊断精度提高20.99%。这表明,最小最大化目标函数有效提高了模型对数据不平衡故障的辨识性能。

为能够更清晰地展示MMOCNN模型对不平衡数据集的特征学习效果,将测试集分别输入数据集 $B$ 训练的CNN和MMOCNN中,使用核主成分分析(kernel principal components analysis,简称KPCA)维数约简算法对模型的FC层输出特征进行维数为2的约简并可视化,结果如图5所示。可以看出:CNN学习的特征比较分散,多数类样本基本能够准确分开,但少数类BL与BS发生了严重混叠,少数类ORL与ORS及IRS发生了严重混叠,如图5(a)所示;MMOCNN学习的特征相对比较聚集,各类之间没有出现严重的混叠现象,仅有少数类BL和ORL中少量样本被错分,如图5(b)所示。这与图4(c,d)的结论基本一致,说明融入最小最大化目标函数可以增强模型在不平衡数据集上的特征学习能力。

为定量分析CNN与MMOCNN特征学习的能力,用测试集分别输入训练的模型中,提取FC层的深度特征,采用式(10)、式(11)来进行评价。 $S_B/S_W$ 和 $(S_B+S_W)/S_W$ 的值越大,表明样本特征的同类距离越小,异类距离越大,学习特征的质量越高。评价的结果如表3所示。

由表3可看出:MMOCNN与CNN在数据集 $A$ 上学习的特征质量相近,说明两个模型在平衡数据集上均能学习较高质量的特征,从而实现故障类别的高精度辨识,这与图4(a,b)的结论一致;但在数

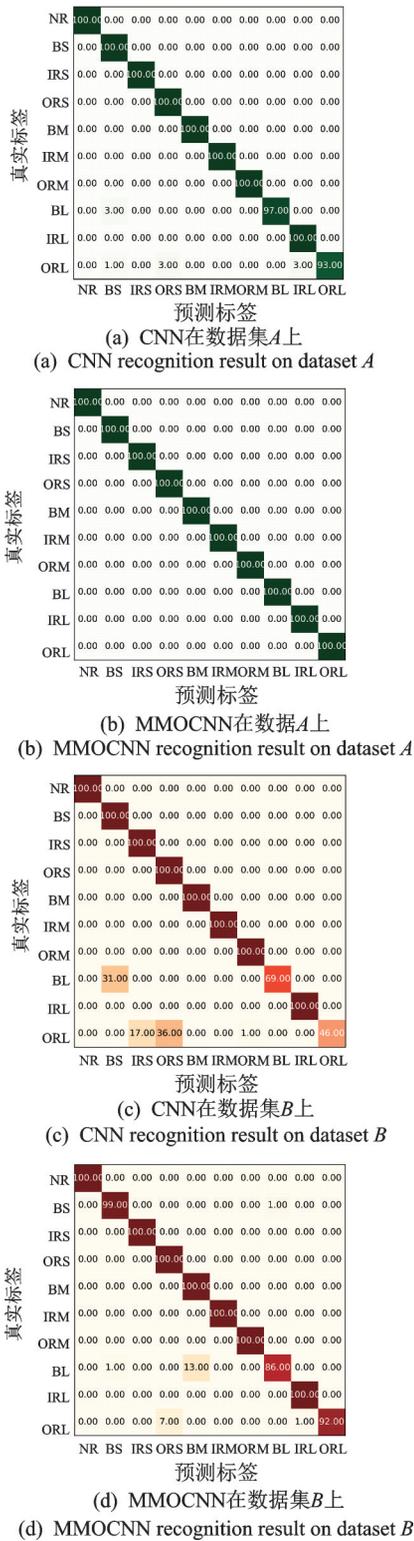


图 4 两模型在测试集上的混淆矩阵

Fig.4 Confusion matrix of two model on test set A and B

据集 B 上, MMOCNN 学习特征的  $S_B/S_w$  和  $(S_B+S_w)/S_w$  的值分别为 55.36 和 56.35, 明显高于 CNN 的 27.48 和 28.48, 这说明 MMOCNN 在不平衡数据集上学习特征的质量更高, 可以更好地实现故障辨识, 这与图 5、图 4(c, d) 的结论一致, 再次说明了最

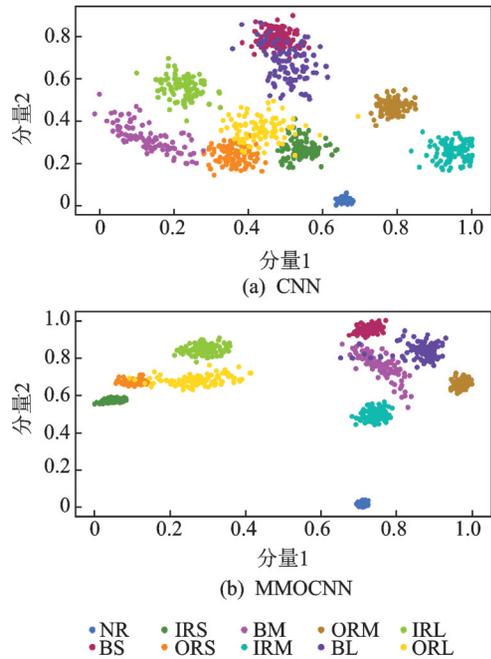


图 5 FC 层提取特征的可视化结果

Fig.5 Visualization of features extracted from FC layer

表 3 特征学习质量的定量评价

Tab.3 Quantitative evaluation of feature learning quality

数据集	方法	评价指标	
		$S_B/S_w$	$(S_B+S_w)/S_w$
A	CNN	30.92	31.93
	MMOCNN	31.09	32.09
B	CNN	27.48	28.48
	MMOCNN	55.36	56.35

小最大化目标函数可以提高模型在不平衡数据集上的特征学习能力。

### 3.3 对比实验情况

为验证 2.3 节所提方法在数据不平衡故障辨识中具有一定优势, 本研究借用表 4 中列出的故障辨识算法进行对比实验。

1) 与手动提取特征的方法进行对比。根据文献[26], 将提取的 24 种特征分别输入到支持向量机 (support vector machine, 简称 SVM) 和反向传播神经网络 (back propagation neural network, 简称 BPNN) 进行数据不平衡故障辨识, 其中 SVM 的核函数使用高斯径向基函数 (radial basis function, 简称 RBF),  $C=30$ ,  $\gamma=0.015$ ; BPNN 结构为 24-100-10, 其激活函数为 ReLU。

2) 与常用的数据不平衡故障诊断方法代价敏感卷积神经网络 (cost sensitive convolutional neural network, 简称 CSCNN) 和 SMOTE+CNN 进行对比。其中: CSCNN 的结构和训练参数与本研究方法一致, 代价矩阵由样本比例确定; 不平衡训练样本

经 SMOTE 采样后变成平衡样本,然后用 CNN 进行故障辨识,CNN 参数与本研究模型保持一致。表 4

总结了本研究方法与 4 种常用辨识算法在表 2 构造的数据集上的识别结果。

表 4 不同方法在不平衡数据集上的辨识结果

Tab.4 Identification results of imbalanced datasets by different methods

数据集	方法	诊断精度										总体
		NR	BS	IRS	ORS	BM	IRM	ORM	BL	IRL	ORL	
A	SVM	100	98.0	100	100	85.0	99.0	100	77.0	100.0	100.0	95.9
	BPNN	100	97.0	100	100	90.0	98.0	100	83.0	98.0	100.0	96.6
	MMOCNN	100	100.0	100	100	100.0	100.0	100	100.0	100.0	100.0	100.0
B	SVM	100	100.0	100	100	73.0	99.0	100	28.0	98.0	100.0	89.8
	BPNN	100	99.0	100	100	79.0	98.0	100	18.0	85.0	100.0	87.9
	SMOTE+CNN	100	97.0	100	100	100.0	100.0	100	71.0	99.0	57.0	92.4
	CSCNN	100	96.0	100	100	100.0	99.0	100	77.0	100.0	84.0	95.6
	MMOCNN	100	99.0	100	100	100.0	100.0	100	86.0	100.0	92.0	97.7

表 4 表明,与 SVM 和 BPNN 方法相比,在平衡数据集 A 上,手工提取的特征分别取得了 95.9% 和 96.6% 的辨识精度,而本研究方法的辨识精度为 100%,具有一定的优势,这是因为手动提取的特征不全面,关键特征有可能遗漏。在不平衡数据集 B 上,手动特征的总体辨识精度分别 89.8% 和 87.9%,而在少数类样本上的辨识精度仅为 75.33% 和 67.67%,这说明手动提取的特征对数据不平衡更加敏感,本研究方法在不平衡数据集上依然保持 97.7% 的辨识精度,具有明显的优势。因此,与手工提取特征相比,图 1 模型更适用于数据不平衡的故障辨识。

与 SMOTE+CNN 和 CSCNN 相比,在不平衡数据集 B 上 3 个模型均能获得 92% 以上的总体辨识精度,这说明 SMOTE 采样和损失函数加权在一定程度上能够缓解数据不平衡对模型性能的影响;但是在少数类样本上,前两个模型的诊断精度分别为 75.67% 和 87%,而本研究的诊断精度为 92.66%,能够得到更高的识别率,说明本研究模型通过在损失函数中融入最小最大化目标函数,模型在数据不平衡故障辨识中对少数类样本取得更好的辨识性能。

### 3.4 模型的泛化性能测试

MMOCNN 模型在公开的轴承数据集上对样本分布不平衡的少数类样本能够取得一个较高的辨识精度,但模型可能存在一定的局限性。为验证本研究 MMOCNN 模型的泛化性能,采用综合故障模拟实验台模拟滚动轴承的运行状态,采集不同运行状态下的振动信号进行辨识,如图 6 所示。实验过程中由加速度传感器采集轴承(NSK6308)在滚动体故障(BF)、内圈故障(IF)、外圈故障(OF)、保持架故障(FF)和正常共 5 种状态下的振动信号,如图 7

所示。轴承转速为 3.2 kr/min,采样频率为 8 kHz,每种运行状态采样 300 组,样本的长度为 1 024 点。用采集的样本构造不平衡数据集,如表 5 所示。

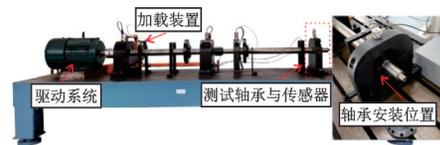


图 6 双跨转子综合故障模拟平台

Fig.6 Double-span rotor comprehensive fault simulation platform

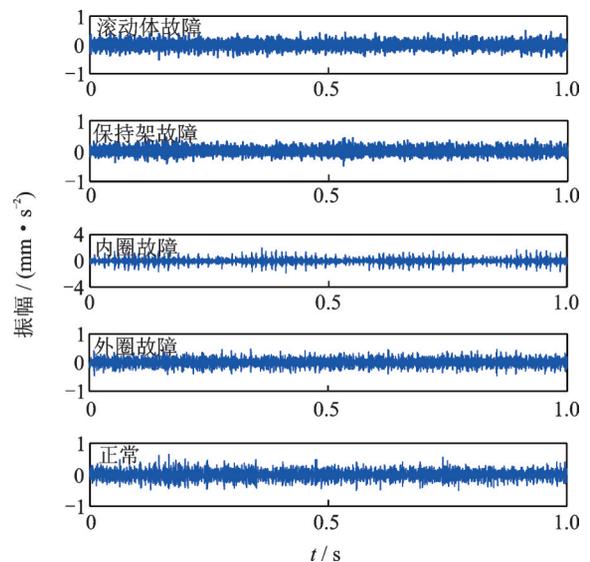


图 7 不同运行状态下的振动信号时域波形

Fig.7 Time-domain waveforms of vibration signals under different operating states

用表 5 的 4 种不平衡训练集,按照图 2 的流程分别训练 MMOCNN 模型,用测试集测试训练模型的辨识精度,结果如图 8 所示。

表 5 构造测试不平衡数据集的描述  
Tab.5 Description of measured datasets

数据集	运行状态					
	NR	BF	FF	IF	OF	
A	训练样本	200	100	60	40	20
	测试样本	100	100	100	100	100
B	训练样本	200	60	40	20	100
	测试样本	100	100	100	100	100
C	训练样本	200	40	20	100	60
	测试样本	100	100	100	100	100
D	训练样本	200	20	100	60	40
	测试样本	100	100	100	100	100

由图 8 可看出:在数据集 A 上,OF 类有 8 个样本被错误识别为 BF 类,IF 类有 1 个样本被错误识别为 OF 类,模型的总体辨识精度为 98.2%,对少数类样本的辨识精度为 92%;在数据集 B 上,IF 类有 16 个样本被错误识别为 OF 类,模型的总体辨识精度为 96.8%,对少数类样本的辨识精度为 84%;在数据集 C 上,模型的总体辨识精度和对少数类样本的辨识精度均为 100%;在数据集 D 上,BF 类有 6 个样本被错误识别为 OF 类,模型的总体辨识精度为 98.8%,对少数类样本的辨识精度为 94%。模型在 4 个不平衡数据集上的总体辨识精度为 98.45%,对少数类样本的平均辨识精度为 92.5%。可见,

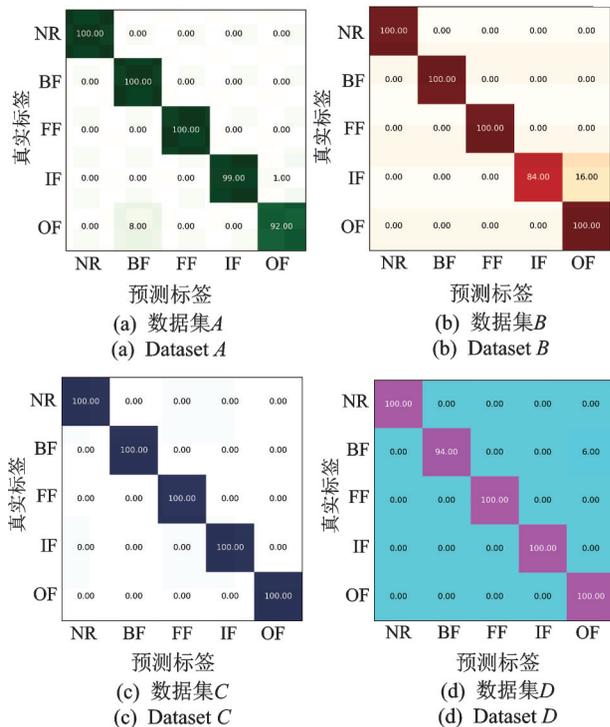


图 8 所提方法在不平衡数据集上的辨识结果

Fig.8 Identification results of the proposed method on imbalanced datasets

MMOCNN 模型在实测模拟轴承数据不平衡故障辨识中对少数类样本也可以获得较高的辨识精度。这表明,针对数据不平衡的故障辨识,本研究方法具有一定的泛化性能。

### 4 结 论

- 1) 笔者建立的 MMOCNN 模型在构造损失函数的同时考虑分类总体精度高、同类距离小和异类距离大,可以对数据不平衡故障中少数类样本实现有效辨识。
- 2) 相比于手动提取特征,特征学习的方法取得了更好的辨识效果。

### 参 考 文 献

[1] LEE J, WU F J, ZHAO W Y, et al. Prognostics and health management design for rotary machinery systems-Reviews, methodology and applications[J]. Mechanical Systems and Signal Processing, 2014, 42(1/2): 314-334.

[2] 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.

[3] 雷亚国, 贾峰, 孔德同, 等. 大数据下机械智能故障诊断的机遇与挑战[J]. 机械工程学报, 2018, 54(5): 94-104. LEI Yaguo, JIA Feng, KONG Detong, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era[J]. Journal of Mechanical Engineering, 2018, 54(5): 94-104. (in Chinese)

[4] 文成林, 吕菲亚. 基于深度学习的故障诊断方法综述 [J]. 电子与信息学报, 2020, 42(1): 234-248. WEN Chenglin, LÜ Feiya. Review on deep learning based fault diagnosis[J]. Journal of Electronics and Information Technology, 2020, 42(1): 234-248. (in Chinese)

[5] ZHAO R, YAN R Q, CHEN Z H, et al. Deep learning and its applications to machine health monitoring [J]. Mechanical Systems and Signal Processing, 2019, 115: 213-237.

[6] 吴耀春, 赵荣珍, 靳伍银. EWT 与加权多邻域粗糙集结合的旋转机械故障特征提取方法[J]. 振动与冲击, 2019, 38(24): 235-242. WU Yaochun, ZHAO Rongzhen, JIN Wuyin. Fault feature extraction of rotating machinery based on EWT and a weighted multi neighborhood rough set[J]. Journal of Vibration and Shock, 2019, 38(24): 235-242. (in Chinese)

[7] 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, 37(6): 1208-1215. WANG Lihua, XIE Yangyang, ZHOU Zixian, et al. Motor fault diagnosis based on convolutional neural net-

- works[J]. Journal of Vibration, Measurement & Diagnosis, 2017,37(6):1208-1215. (in Chinese)
- [8] LU C, WANG Z Y, QIN W L, et al. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification [J]. Signal Processing, 2017,130: 377-388.
- [9] 陈仁祥, 黄鑫, 杨黎霞, 等. 基于卷积神经网络和离散小波变换的滚动轴承故障诊断[J]. 振动工程学报, 2018,31(5): 883-891.  
CHEN Renxiang, HUANG Xin, YANG Lixia, et al. Rolling bearing fault identification based on convolution neural network and discrete wavelet transform[J]. Journal of Vibration Engineering, 2018, 31(5): 883-891. (in Chinese)
- [10] JANSSENS O, SLAVKOVIKJ V, VERVISCH B, et al. Convolutional neural network based fault detection for rotating machinery [J]. Journal of Sound and Vibration, 2016,377(377): 331-345.
- [11] JING L Y, ZHAO M, LI P, et al. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox[J]. Measurement, 2017,111: 1-10.
- [12] WEN L, LI X Y, GAO L, et al. A new convolutional neural network-based data-driven fault diagnosis method [J]. IEEE Transactions on Industrial Electronics, 2018,65(7): 5990-5998.
- [13] ZHANG W, LIC H, PENG G L, et al. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load [J]. Mechanical Systems and Signal Processing, 2018,100(1): 439-453.
- [14] 董勋, 郭亮, 高宏力, 等. 代价敏感卷积神经网络: 一种机械故障数据不平衡分类方法[J]. 仪器仪表学报, 2019,40(12): 205-213.  
DONG Xun, GUO Liang, GAO Hongli, et al. Cost sensitive convolutional neural network: a classification method for imbalanced data of mechanical fault[J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 205-213. (in Chinese)
- [15] 李艳霞, 柴毅, 胡友强, 等. 不平衡数据分类方法综述 [J]. 控制与决策, 2019,34(4): 4-19.  
LI Yanxia, CHAI Yi, HU Youqiang, et al. Review of imbalanced data classification methods [J]. Control and Decision, 2019,34(4): 4-19. (in Chinese)
- [16] 陈旭, 刘鹏鹤, 孙毓忠, 等. 面向不均衡医学数据集的疾病预测模型研究 [J]. 计算机学报, 2019,42(3): 596-609.  
CHEN Xu, LIU Penghe, SUN Yuzhong, et al. Research on disease prediction models based on imbalanced medical data sets [J]. Chinese Journal of Computers, 2019,42(3): 596-609. (in Chinese)
- [17] 楼晓俊, 孙雨轩, 刘海涛. 聚类边界过采样不平衡数据分类方法 [J]. 浙江大学学报(工学版), 2013, 47(6): 944-950.  
LOU Xiaojun, SUN Yuxuan, LIU Haitao. Clustering boundary over-sampling classification method for imbalanced data sets [J]. Journal of Zhejiang University (Engineering Science), 2013, 47(6): 944-950. (in Chinese)
- [18] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: synthetic minority over-sampling technique [J]. Journal of Artificial Intelligence Research, 2011, 16(1): 321-357.
- [19] 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, 2018, 48: 34-50.
- [20] WU Z Y, LIN W F, JI Y, et al. An integrated ensemble learning model for imbalanced fault diagnostics and prognostics [J]. IEEE Access, 2018, 6: 8394-8402.
- [21] JIA F, LEI Y G, LU N, 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.
- [22] WING W Y, ZENG G J, ZHANG J J, et al. Dual autoencoders features for imbalance classification problem [J]. Pattern Recognition, 2016,60: 875-889.
- [23] SHI W W, GONG Y H, TAO X Y, et al. Improving CNN performance accuracies with min-max objective [J]. IEEE Transactions on Neural Networks, 2018, 29(7): 2872-2885.
- [24] WONG W K, SUN M M. Deep learning regularized fisher mappings [J]. IEEE Transactions on Neural Networks, 2011,22(10): 1668-1675.
- [25] Case Western Reserve University. Bearing data center website [EB/OL]. [2019-01-15]. [https://csegroups.case.edu/Bearing data center/home](https://csegroups.case.edu/Bearing%20data%20center/home).
- [26] QU J X, ZHANG Z S, GONG T. A novel intelligent method for mechanical fault diagnosis based on dual-tree complex wavelet packet transform and multiple classifier fusion [J]. Neurocomputing, 2016,171: 837-853.



**第一作者简介:**吴耀春,男,1981年2月生,博士生、讲师。主要研究方向为机械状态监测与故障诊断、旋转机械智能故障诊断。曾发表《EWT与加权多邻域粗糙集结合的旋转机械故障特征提取方法》(《振动与冲击》2019年第38卷)等论文。

E-mail:wuyaochun2004@sina.com