

基于 MCNN-BiGRU-Attention 的轴承故障诊断^①



陈悦然, 牟 莉

(西安工程大学 计算机科学学院, 西安 710048)

通信作者: 陈悦然, E-mail: 386459090@qq.com

摘 要: 传统的故障诊断方法在对轴承故障进行判定时, 存在特征提取不充分, 时序特性运用不完全且计算较为复杂的问题. 对于此问题, 本文提出一种基于多尺度卷积神经网络 (multi-scale convolutional neural network, MCNN)、双向门控循环单元 (bidirectional gated recurrent unit, BiGRU) 和注意力机制 (attention) 的组合诊断方法. 该方法首先使用 MCNN 对信号数据进行多尺度特征提取, 在空间层面上, 实现了对特征的进一步提炼. 其次使用了 BiGRU 网络, 在时间层面上, 从正反两个方向获取时序关系. 接下来引入注意力机制, 忽略一些与结果不相关的信息并且降低信息丢失的风险以提高精度, 在经过全连接层创建映射后, 最后通过 *Softmax* 分类方法完成轴承故障诊断. 本文通过实验, 与 LSTM 模型、GRU 模型、SVM 模型、CNN-BiGRU 等模型进行对比, 实验结果表明, 本文提出的模型相比以往提出的模型准确率更高, 单一工况下的故障诊断准确率达到 98.1%, 多工况条件下平均准确率达到 97.8%.
关键词: 轴承; 故障诊断; 多尺度卷积; 注意力机制; 双向门控循环单元

引用格式: 陈悦然, 牟莉. 基于 MCNN-BiGRU-Attention 的轴承故障诊断. 计算机系统应用, 2023, 32(9): 125-131. <http://www.c-s-a.org.cn/1003-3254/9251.html>

Bearing Fault Diagnosis Based on MCNN-BiGRU-Attention

CHEN Yue-Ran, MU Li

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract: When the traditional fault diagnosis method is used to judge the bearing fault, there are some problems such as insufficient feature extraction, incomplete use of time sequence characteristics, and complicated calculation. In order to solve these problems, a combined diagnosis method based on a multi-scale convolutional neural network (MCNN), bidirectional gated recurrent unit (BiGRU), and attention mechanism is proposed in this study. Firstly, MCNN is used to extract multi-scale features from signal data, which realizes further extraction of features in terms of spatial level. Secondly, the BiGRU network is used to obtain sequential relations from positive and negative directions in terms of time level. Next, an attention mechanism is introduced to ignore some information that is not relevant to the results and reduce the risk of information loss to improve accuracy. After mapping is created through the full connection layer, the bearing fault diagnosis is finally completed through the *Softmax* classification method. In this study, LSTM model, GRU model, SVM model, CNN-BiGRU model, and other models are compared through experiments, and the experimental results show that the proposed model has higher accuracy than the previous models. The accuracy of fault diagnosis in a single working condition reaches 98.1%, and the average accuracy of fault diagnosis in multiple working conditions reaches 97.8%.
Key words: bearing; fault diagnosis; multiscale convolution; attention mechanism; bidirectional gated recurrent unit (BiGRU)

① 收稿时间: 2023-03-10; 修改时间: 2023-04-10; 采用时间: 2023-04-28; csa 在线出版时间: 2023-07-21
CNKI 网络首发时间: 2023-07-21

推进新型工业化,加快建设数字强国,成为我国实现百年奋斗目标的要领之一.工业化的稳步推进,离不开各种工业机械设备的加持,轴承作为各种旋转机械设备中的关键部件,一旦出现异常,轻则损坏设备,重则造成安全事故.而轴承故障诊断可以快速判断出轴承的故障类型,这对于轴承的维护及保养具有重要意义.目前,轴承的故障诊断可以分为3种,一种是传统的基于信号特征的诊断方法,如基于时域分析^[1]、频域分析^[2]以及时频域分析^[3],然而单一的基于信号特征的方法人为干预较强,某些参数可能会受到主观因素的影响.因此研究人员开始将机器学习与时频分析等相结合.He等人^[4]使用时频分析中的小波包能量谱方法,提取信号中的特征,引入支持向量机(SVM),诊断出轴承故障.谢小正等人^[5]等利用灰狼算法优化SVM参数,提出一种基于小波包能量熵的轴承诊断方法.虽然以上方法效果尚可,但本质上都基于机器学习.在数据量过大时,求解效率很低,导致预测效果欠佳.随着深度学习的兴起,研究人员开始将深度学习与轴承故障诊断结合.刘颖等人^[6]利用小波包分解对信号特征进行提取,再输入至卷积神经网络(CNN),最终完成故障诊断.陈宇航等人^[7]使用快速傅里叶变换得到频域特征,提出一种基于CNN的故障诊断方法.然而,采集到的轴承振动信号带有时序关系,仅用CNN进行训练,难以得到隐藏的时序关系.循环神经网络(RNN)由于其上一步的输入会对下一步的结果产生影响,很好地体现了数据中的时序关系.庄夏^[8]采用了RNN检测电机轴承故障.RNN虽然解决了获取时序关系的问题,但是如果时间序列过长,RNN网络就会出现梯度爆炸或梯度消失的现象.长短期记忆网络(LSTM)作为RNN的变体,很大程度上减轻了梯度爆炸或消失的现象.唐赛等人^[9]将得到的特征输入至LSTM网络训练,完成故障诊断.但由于LSTM特别的三门机制,计算量过大.循环门控单元(GRU)优化LSTM,在结构上做出了极大简化.张伟等人^[10]提出一种基于GRU的轴承故障诊断方法,预测精度比以往得到提高.BiGRU是近几年新提出的一种模型^[11],不但可以获取时间间的关系,还可以做到正向获取与反向获取,同时考虑前后的信息,使得到的时序特征关系更加丰富.石静雯等人^[12]将原始振动信号直接输入至BiGRU网络,诊断结果准确率良好.但是没有考虑将原始信号先进行处理再训练,蒋小平等人^[13]在工况识别时使用了MCNN对数据进

行了处理,结果表明,使用MCNN提取特征比CNN准确率更高.聂良鹏等人^[14]、安文杰等人^[15]将MCNN与BiGRU结合,证明了组合模型更有优势.但是上述组合模型没有使用attention机制,使得实验结果精度还有进一步上升空间.

基于上述,为了使最终实验结果的精度进一步提高,轴承故障的预测更加有效,结合现有的各种单一模型优势,本文提出一种MCNN-BiGRU-Attention组合模型对轴承故障诊断展开研究与验证.该模型的优点在于,能够充分利用数据的多种特征和上下文信息,同时通过注意力机制加强模型对关键信息的关注,从而提高了模型的性能.具体来说,将振动信号用MCNN进行更丰富的特征提取,即使用不同尺寸的卷积核提取特征并融合,其次输入至故障诊断部分,使用BiGRU神经网络进行训练,从前后两个方向对融合好的特征进行更深层次的信息挖掘,通过attention机制,给予故障特征更高的权重,确保重要信息,实验表明MCNN-BiGRU-Attention模型诊断的拟合程度更好,准确率更高.

1 MCNN-BiGRU-Attention 模型

1.1 卷积神经网络

卷积神经网络是一种由感知机演化而来的前向传播的深度网络,相比于支持向量机,具备更好的数据适应性以及特征提取能力^[16].卷积神经网络的隐藏层通常由若干卷积层与池化层构成.其中,卷积层具有稀疏交互与权值共享等属性,可以有效降低训练所需的参数要求,防止因参数过多造成的过拟合.卷积的运算过程如下:

$$y_j^l = \sigma \left(\sum_{i \in R_j} (x_i^{l-1} k_{ij}^l) + b_j^l \right) \quad (1)$$

其中, y_j^l 为第 l 层的输出, σ 为激活函数, b_j^l 为第 l 层的偏置, k_{ij}^l 为权重, x_i^{l-1} 为第 $l-1$ 层的输出, R_j 为输入的特征集合.

通常情况下卷积层的后面会衔接一层池化层,以降低特征的数量,减少计算量与内存消耗.常见的池化操作有平均池化与最大池化,表达式分别为:

$$v^{l(i,j)} = \max_{H(j-1)-1 < n < jH} \{x^{l(i,n)}\} \quad (2)$$

$$v^{l(i,j)} = \frac{1}{H} \sum_{H(j-1)-1}^{jH} x^{l(i,n)} \quad (3)$$

其中, $l(i, n)$ 表示第 l 层的 i 通道第 n 神经元的激活值, H 为滑动窗口大小, $v^{(i,j)}$ 表示当前第 l 层第 j 个神经元池化后的结果.

1.2 门控循环单元

门控循环单元 (gated recurrent unit, GRU) 是在 LSTM 的基础上演化而来^[17], 改进了 LSTM 的结构, 使得 GRU 模型从复杂的三门结构, 变为只有更新与重置两个门, 重置门用来获取长期关系, 更新门则可以获取短期关系. 其不但继承了 RNN 可以处理时间序列的能力, 还延续了 LSTM 能最大程度上减轻序列过长时存在的梯度爆炸或梯度消失问题, 同时对新样本的适应能力更强. GRU 的结构如图 1 所示.

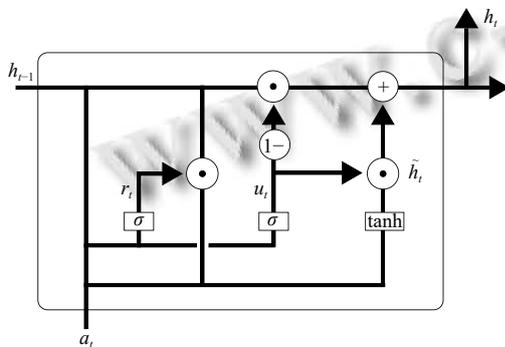


图 1 GRU 神经网络单元结构图

对于当前时间步的输入和上一个时间步的输出, GRU 的运算步骤如下.

$$u_t = f(W_u \cdot [a_t, h_{t-1}]) \quad (4)$$

$$r_t = f(W_r \cdot [a_t, h_{t-1}]) \quad (5)$$

$$\tilde{h}_t = \tanh(W \cdot [a_t, r_t \odot h_{t-1}]) \quad (6)$$

$$h_t = [(1 - u_t) \odot h_{t-1}] \oplus [u_t \odot \tilde{h}_t] \quad (7)$$

其中, r_t 和 u_t 分别为当前时刻重置门输出结果和更新门输出结果, \tilde{h}_t 为候选隐状态, 用来描述当前时刻的输入状态和更新门的状态, h_t 为隐状态, 如果 u_t 的值接近 0, 则隐状态会更加关注 t 时刻之前的信息, 反之则接近候选隐状态, f 为激活函数, W 为权重构成的矩阵.

GRU 是网络的传播方向是从左到右的, 即正向传播. 而为了能够更准确地获取信号之间潜在的时序关系, 还要做到反向捕获信息, 因此引入双向门控循环单元, 即 BiGRU (bidirectional gated recurrent unit).

1.3 Attention 机制

Attention 机制的思想是模仿了人类在看到某处画

面时, 会最先注意到画面中重要的信息, 而一些非重要的信息则会被忽视. 这一机制使得当神经网络发现输入数据的关键信息后, 通过学习, 在后继的预测阶段对其予以重点关注^[18]. Attention 机制使用在图像处理^[19], 自然语言处理^[20] 等领域得到了较好的效果, 本文中使用的注意力机制, 根据各个特征对于预测结果的重要程度, 给予隐藏层输出对应的权重. 能对决定最终结果走向的因素, 应当给予更高的权重, 以此来最大程度利用重要信息, 从而提高预测结果精度.

1.4 MCNN-BiGRU-Attention 故障诊断组合模型

本文针对以往的故障诊断模型提取特征不充分以及对重点信息关注不足的问题, 提出了 MCNN-BiGRU-Attention 故障诊断模型, 该模型可分为 4 部分, 其结构如图 2 所示.

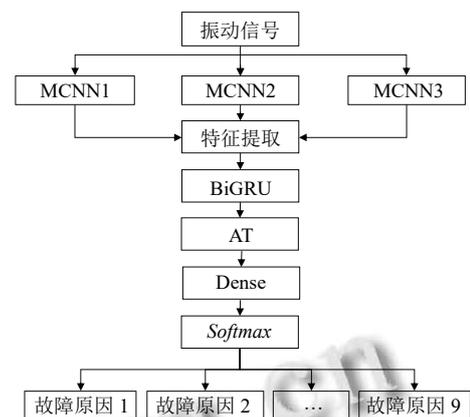


图 2 MCNN-BiGRU-Attention 结构示意图

(1) MCNN: 多尺度卷积神经网络 (multi-scale convolution neural network, MCNN) 是 CNN 的变体, 拥有比 CNN 更加强大的特征提取能力. 其使用不同尺寸的卷积核对特征进行提取. 本文使用若干较大型、中型以及较小型的卷积核, 这样做的好处是, 不同尺寸卷积核具有不同大小的感受野, 从而可以提取到低频率以及高频的特征, 将这些特征叠加起来, 以形成多尺度的特征. 使用多尺度的特征进行训练, 其结果要优于单一卷积神经网络提取的特征, 池化层采用平均池化, 减小参数量.

(2) BiGRU: 将堆叠后的特征输入至 BiGRU 网络, 由于 BiGRU 是由正向及反向的两个 GRU 模型组合, BiGRU 的结构如图 3 所示.

由于输入的是整个时间序列, 所以在训练时, 某个

时刻的特征既依赖于之前的信息,又有该时刻未来的信息作为依据,计算公式如下:

$$h_1^{(t)} = \delta(\vec{W} \cdot h_1^{(t-1)} + \vec{U} \cdot x^{(t)}) \quad (8)$$

$$h_2^{(t)} = \delta(\overleftarrow{W} \cdot h_2^{(t+1)} + \overleftarrow{U} \cdot x^{(t)}) \quad (9)$$

$$H^{(t)} = [h_1^{(t)}, h_2^{(t)}] \quad (10)$$

其中, $h_1^{(t)}$ 为前向传播的隐藏状态, $h_2^{(t)}$ 为后向传播的隐藏状态, $h_1^{(t-1)}$ 为相对于某时刻的前一个时刻的隐藏状态, $h_2^{(t+1)}$ 为相对于某时刻的后一个时刻的隐藏状态, δ 为激活函数, $x^{(t)}$ 为 t 时刻的输入, $H^{(t)}$ 为 t 时刻的输出, U, W 为权重矩阵, 箭头为传播方向.

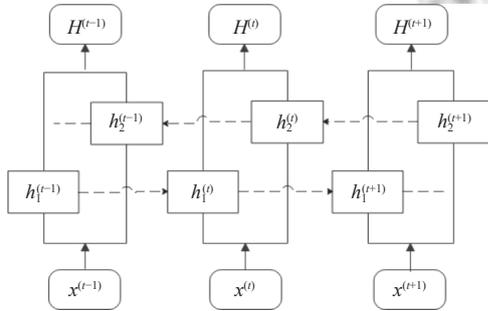


图3 BiGRU神经网络结构图

因此可以从前后两个方向挖掘潜在的信息,得到的结果应为 BiGRU 训练后的双向时序序列:

$$O^{(s)} = [O^{(1)}, \dots, O^{(t)}, \dots, O^{(T)}]^T \quad (11)$$

其中, $O^{(t)}$ 为 t 时刻 BiGRU 的输出结果, T 为总时间步长.

(3) Attention: 考虑到不是所有 BiGRU 输出的特征向量都对轴承故障诊断起决定性作用,对不同 BiGRU 隐状态进行归一化处理,并赋予不同的权重,以此来达到关注重要信息的效果.其工作原理如图4所示. Attention 机制的表达式如下:

$$w_T = \text{Softmax}(k_T h_T + b_T) \quad (12)$$

$$S = \sum_{i=1}^T w_i \cdot h_i \quad (13)$$

其中, h_i 为某时刻 BiGRU 网络的隐藏层输出, b 和 k 分别为偏置与权重, w_i 为权重的分配, $w_1 + w_2 + \dots + w_T = 1$, T 为总的时间步长, S 为最终的表征向量.

(4) 故障分类层: 包括全连接与 Softmax 层, 全连接

层每个神经元与上层的神经元相连,将提取到的特征进行整合, Softmax 层进行故障分类,得到的结果为一个 0 到 1 之间的数,最终与提前设置好的故障值对比,比重较大的即可认定为该类型的故障,从而完成故障诊断, Softmax 公式如下:

$$p = \frac{e^{v_i}}{\sum_k^n e^{v_k}} \quad (14)$$

其中, k 为具体某个分类, n 为分类个数, v 为分类的值.

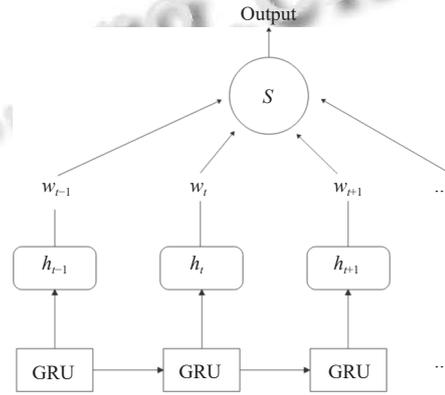


图4 Attention 机制示意图

2 实验分析

2.1 实验环境与数据

本文的实验数据选自美国凯斯西储大学 (CWRU) 公开轴承数据集^[21],该数据集在轴承故障研究领域内被公认为最具价值的数据集.数据的采集由频率为 12 kHz 的驱动端传感器完成,轴承的故障植入采用了电火花加工 (EDM) 技术,根据电火花蚀刻位置的不同,将故障分为了内圈故障,外圈故障以及滚动体故障每种故障的损坏范围不同,分别为 0.007 inch, 0.014 inch, 0.021 inch, 即轻度,中度,重度,考虑到实际情况的工况较复杂,又分为了 3 种工况.如表 1 所示.数据处理部分,本文使用滑动窗口法^[22],对每种状态的数据进行增广,如图 5 所示,滑动窗口范围内的数据即为每个样本包含的数据,设为 2 048 个,总的训练集为 5 000 个样本,测试集为 1 000 个样本.轴承的故障状态与正常状态共有 10 种,分别用数字 0-9 表示.实验所用的框架为 PyTorch,运行环境的 GPU 为 RTX3050Ti.

2.2 参数选择

MCNN-BiGRU-Attention 模型的多尺度特征提取部分,MCNN1 采用的卷积核大小为 16×1 、 8×1 ,步长

为2, 个数分别为40和20; MCCN2采用的卷积核大小为4×1, 步长为1, 个数分别为40和30; MCNN3采用的卷积核尺寸为4×1, 卷积核个数为20, 激活函数选用ReLU^[23]. 时序关系的提取部分, BiGRU网络的激活函数选用Sigmoid, 梯度下降选用Adam算法^[24].

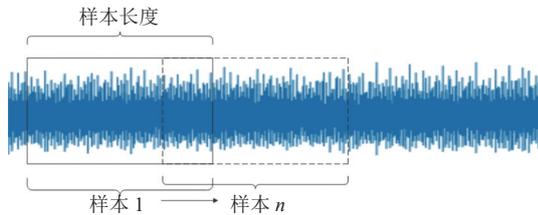


图5 采样示意图

表1 样本数据

故障类别	故障程度 (inch)	故障位置	训练样本数量	测试样本数量	样本长度	转速 (rpm)
0	0.007	内圈	500	100	2048	1750-1797
1	0.007	外圈	500	100	2048	1750-1797
2	0.007	滚动体	500	100	2048	1750-1797
3	0.014	内圈	500	100	2048	1750-1797
4	0.014	外圈	500	100	2048	1750-1797
5	0.014	滚动体	500	100	2048	1750-1797
6	0.021	内圈	500	100	2048	1750-1797
7	0.021	外圈	500	100	2048	1750-1797
8	0.021	滚动体	500	100	2048	1750-1797
9	正常	—	500	100	2048	1750-1797

选用 dropout 来避免 BiGRU 训练过拟合, 学习率为 0.001. 故障分类部分, 选用的激活函数为 Softmax.

2.3 单一负载下的轴承故障诊断

单一工况下的轴承故障诊断选用 0 HP 时的数据, 由于实验所用到的训练集和测试集的分配方式为随机分配, 为了减少随机分配带来的偶发性因素影响, 实验将采用交叉验证的方法, 将得到的 600 个数据样本分为 6 份, 每份 100 个样本. 第 1 次实验, 取第 1 份样本作为测试集, 取后 5 份样本作为训练集进行实验, 第 2 次实验, 取第 2 份样本作为测试集, 取该样本外的其他 5 份样本作为训练集, 以此类推, 最终结果为 6 次实验的平均结果. 为了避免过多的迭代次数, 更快速地得到结果, 在 MCNN-BiGRU-Attention 网络中添加条件, 当迭代次数小于 20 次的同时, 损失值小于 0.2 就可以终止迭代, 从而快速做出诊断. 实验迭代 60 次的准确率曲线与 loss 值曲线如图 6 所示.

对比训练与测试两个曲线可以得出, 模型在第

40 次迭代的时候, 训练集的准确率为 97.46%, 测试集的准确率为 97.12%, 两条曲线逐渐收敛, 模型逐渐趋于稳定. 为了验证本文提出模型的有效性与其优越性, 本文将该模型与 LSTM、GRU、BiGRU、CNN-BiGRU、MCNN-BiGRU 等其他基于深度学习预测模型进行轴承故障诊断预测, 进行对比, 如图 7 和图 8 所示. 实验结果记录如表 2 所示.

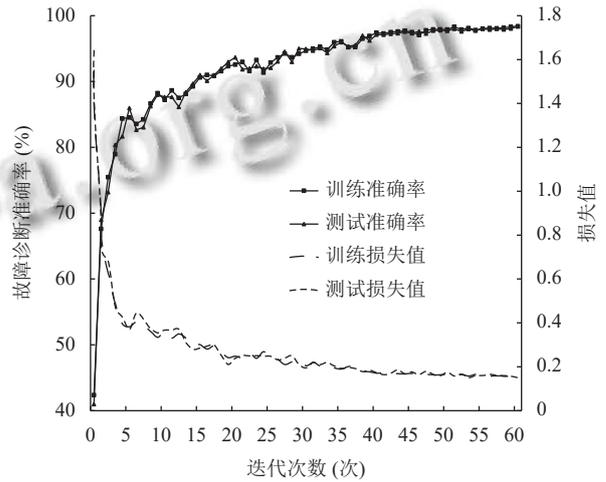


图6 MCNN-BiGRU-Attention 的准确率与损失值

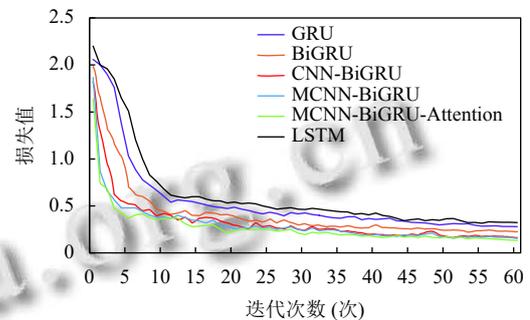


图7 一定迭代次数下不同模型损失值对比

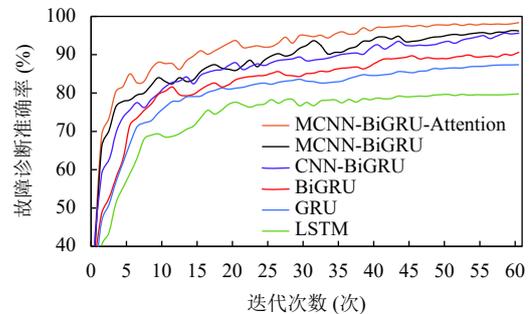


图8 一定迭代次数下不同模型准确率对比

LSTM 与 GRU 的单一模型由结构相较于其他模型较为简单, 因此训练所花费的时间较短, BiGRU 模型

是由两个异向 GRU 模型组合,能进行双向学习,与 GRU 模型相比,训练效果更好,故障识别率更高.带有 CNN 结构的 CNN-BiGRU 模型,相比于单一网络模型,故障识别效果更好,这是由于通过 CNN 在空间层面提取特征,通过 BiGRU 在时间层面提取特征,两者组合的特征提取效果要好于单一模型,但是训练时间也会因此增加.与 CNN-BiGRU 相比,MCNN-BiGRU 由于采用多种不同卷积核提取特征,保证了特征的全面性与丰富性,因此诊断精度更高.本文提出的 MCNN-BiGRU-Attention 模型相比上述模型,训练时间与其他组合模型相差无几,但准确率最高,标准差最低,具有良好的健壮性,证明了引入 attention 机制后,本文提出的模型在轴承故障诊断方面取得的最优的效果.

表 2 不同神经网络诊断准确率对比

诊断模型	准确率 (%)	标准差
LSTM	79.5	2.78
GRU	86.3	1.54
BiGRU	90.6	0.96
CNN-BiGRU	95.6	0.62
MCNN-BiGRU	96.2	0.25
MCNN-BiGRU-Attention	98.1	0.18

2.4 不同负载下的轴承故障诊断

在实际生产生活中,轴承的负载并不是一成不变的,不同情况下的轴承载荷有所不同,导致了不同载荷下的故障信号也有差异,如图 9 所示.

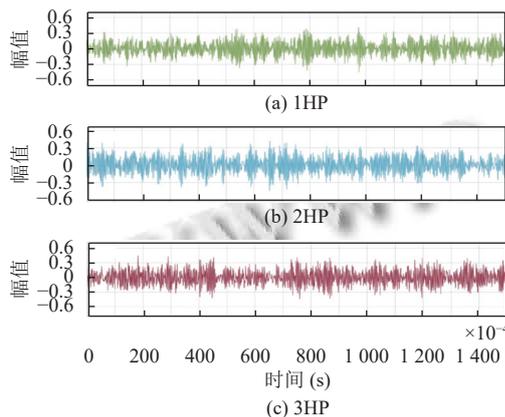


图 9 不同工况下的滚动体 0.007 inches 故障信号

为了验证方法的鲁棒性,因此还需考虑不同工况下的诊断效果.分别选用 1HP, 2HP, 3HP 这 3 种工况中的一种数据集作为训练集,另外两种作为测试集进行实验,选用机器学习中的 SVM 方法和基于深度学习

模型的 GRU, BiGRU, CNN-BiGRU, MCNN-BiGRU 方法作为对比组与本文方法进行对比.实验结果如图 10 所示,其中,“1HP/2HP”为使用 1HP 负载的数据集训练,使用 2HP 负载的数据集测试. AVG 为平均的诊断准确率.从图 10 中得知,基于深度学习的方法的诊断精度要远大于机器学习中的 SVM 方法,组合网络的诊断精度要高于单一网络的预测精度,准确率达到 90% 以上.当工况为 2HP/3HP 时,在其他模型准确率普遍偏低的情况下,本文提出的模型表现最稳定,诊断准确率最优.与 CNN-BiGRU 模型相比,加入 attention 机制能对够重要信息进行优化,同时体现出了使用 MCNN 后提取到的特征更加充分,本文模型的平均准确率接近 98%,提升了 3%.

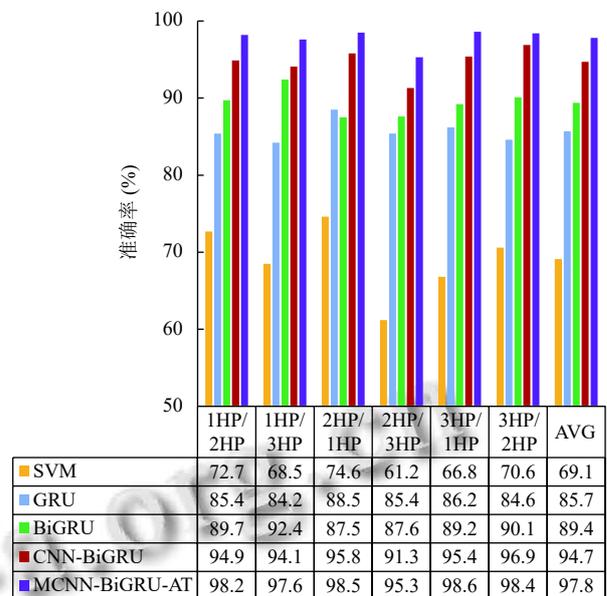


图 10 不同负载下的各模型准确率对比

3 结论

本文提出了 MCNN-BiGRU-Attention 模型进行故障诊断,该方法首先利用 MCNN 进行多尺度特征提取,深度挖掘原始数据信号中的潜在信息,更加充分地提取了特征.其次使用 BiGRU 进行正反两个方向的训练,最大化的获信号中的时序关系,并且引入 attention 机制,尽可能在训练时保留重要的信息,最后通过 Softmax 实现轴承故障的诊断.克服了以往深度学习特征提取与时序关系利用不充分的问题.单一负载下的实验结果表明,本文提出的 MCNN-BiGRU-Attention 模型的

故障诊断精度与其他传统模型相比最高,达到了98.1%;其次,通过不同负载下的实验验证,MCNN-BiGRU-Attention模型的诊断精度都优于其他4种模型,其平均精度也达到97.8%,对比其他故障诊断模型,本文模型在面对不同的负载时表现出更高的准确性与健壮性。

参考文献

- 1 Zhang X, Wan ST, He YL, *et al.* Bearing fault diagnosis based on iterative 1.5-dimensional spectral kurtosis. *IEEE Access*, 2020, 8: 174233–174243.
- 2 王旭峰. 基于倒频谱分析法的滚动轴承故障诊断研究 [硕士学位论文]. 昆明: 昆明理工大学, 2010.
- 3 Li D. Research of fault diagnosis of mine rolling bearing based on time-frequency analysis. *Journal of Physics: Conference Series*, 2023, 2459(1): 012081. [doi: [10.1088/1742-6596/2459/1/012081](https://doi.org/10.1088/1742-6596/2459/1/012081)]
- 4 He C, Wu T, Gu RW, *et al.* Bearing fault diagnosis based on wavelet packet energy spectrum and SVM. *Journal of Physics: Conference Series*, 2020, 1684(1): 012135. [doi: [10.1088/1742-6596/1684/1/012135](https://doi.org/10.1088/1742-6596/1684/1/012135)]
- 5 谢小正, 王晋, 赵荣珍, 等. 基于小波包能量熵和GWO-SVM的滚动轴承故障诊断. *兰州理工大学学报*, 2022, 48(5): 59–64.
- 6 刘颖, 陶建峰, 黄武涛, 等. 小波包能量与CNN相结合的滚动轴承故障诊断方法. *机械设计与制造*, 2021, (11): 127–131. [doi: [10.3969/j.issn.1001-3997.2021.11.029](https://doi.org/10.3969/j.issn.1001-3997.2021.11.029)]
- 7 陈宇航, 李正平, 肖雷. 基于FFT-1D-CNN的细纱机罗拉轴承故障诊断. *棉纺织技术*, 2023, 51(1): 16–21. [doi: [10.3969/j.issn.1001-7415.2023.01.004](https://doi.org/10.3969/j.issn.1001-7415.2023.01.004)]
- 8 庄夏. 基于DWT和RNN的无刷直流电动机轴承故障检测方法. *微特电机*, 2017, 45(6): 17–21, 26. [doi: [10.3969/j.issn.1004-7018.2017.06.005](https://doi.org/10.3969/j.issn.1004-7018.2017.06.005)]
- 9 唐赛, 何苻兮, 张家悦, 等. 基于长短期记忆网络的轴承故障识别. *汽车工程学报*, 2018, 8(4): 297–303. [doi: [10.3969/j.issn.2095-1469.2018.04.09](https://doi.org/10.3969/j.issn.2095-1469.2018.04.09)]
- 10 张伟, 张广帅, 王连彪. 基于CNN-GRU网络的轴承故障检测算法. *工业仪表与自动化装置*, 2021, (6): 88–91. [doi: [10.3969/j.issn.1000-0682.2021.06.019](https://doi.org/10.3969/j.issn.1000-0682.2021.06.019)]
- 11 Lei DX, Liu HJ, Le HJ, *et al.* Ionospheric TEC prediction base on attentional BiGRU. *Atmosphere*, 2022, 13(7): 1039. [doi: [10.3390/atmos13071039](https://doi.org/10.3390/atmos13071039)]
- 12 石静雯, 侯立群. 基于双向门控循环单元网络的滚动轴承故障诊断. *电力科学与工程*, 2021, 37(10): 64–70. [doi: [10.3969/j.issn.1672-0792.2021.10.008](https://doi.org/10.3969/j.issn.1672-0792.2021.10.008)]
- 13 蒋小平, 刘俊威, 王乐乐, 等. 基于多尺度卷积神经网络和LBP算法的浮选工况识别. *矿业科学学报*, 2023, 8(2): 202–212. [doi: [10.19606/j.cnki.jmst.2023.02.007](https://doi.org/10.19606/j.cnki.jmst.2023.02.007)]
- 14 聂良鹏, 权丽君, 吴庭芳, 等. BMBQA: 融合MCNN和BiGRU的蛋白质模型质量评估算法. *小型微型计算机系统*, 2022, 43(7): 1419–1425. [doi: [10.20009/j.cnki.21-1106/TP.2020-1083](https://doi.org/10.20009/j.cnki.21-1106/TP.2020-1083)]
- 15 安文杰, 陈长征, 田淼, 等. 基于MSCNNSA-BiGRU的变工况风电机组滚动轴承故障诊断研究. *机电工程*, 2022, 39(8): 1096–1103. [doi: [10.3969/j.issn.1001-4551.2022.08.010](https://doi.org/10.3969/j.issn.1001-4551.2022.08.010)]
- 16 李辰, 李建勋. 卷积神经网络的正交性特征提取方法及其应用. *上海交通大学学报*, 2021, 55(10): 1320–1329. [doi: [10.16183/j.cnki.jsjtu.2020.276](https://doi.org/10.16183/j.cnki.jsjtu.2020.276)]
- 17 Cho K, van Merriënboer B, Gulcehre C, *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv:1406.1078, 2014.
- 18 石磊, 王毅, 成颖, 等. 自然语言处理中的注意力机制研究综述. *数据分析与知识发现*, 2020, 4(5): 1–14.
- 19 Lee H, Cho S. Locally adaptive channel attention-based network for denoising images. *IEEE Access*, 2020, 8: 34686–34695. [doi: [10.1109/ACCESS.2020.2974001](https://doi.org/10.1109/ACCESS.2020.2974001)]
- 20 Shi NW, Wang HZ, Cheng YQ. AT-CRF: A chinese reading comprehension algorithm based on attention mechanism and conditional random fields. *Applied Sciences*, 2022, 12(20): 10459. [doi: [10.3390/app122010459](https://doi.org/10.3390/app122010459)]
- 21 Smith WA, Randall RB. Rolling element bearing diagnostics using the case Western Reserve University data: A benchmark study. *Mechanical Systems and Signal Processing*, 2015, 64–65: 100–131.
- 22 刘洋. 基于卷积神经网络的风电机组滚动轴承故障诊断研究 [硕士学位论文]. 南京: 东南大学, 2021.
- 23 汪丁, 黄葵, 朱兴动, 等. 基于MCNN-GRU的舰面目标碰撞预警方法. *兵工自动化*, 2022, 41(8): 52–57, 80.
- 24 张玺君, 尚继洋, 余光杰, 等. 基于注意力的多尺度卷积神经网络轴承故障诊断. *吉林大学学报(工学版)*. <http://kns.cnki.net/kcms/detail/22.1341.T.20230407.1503.006.html>. [2023-04-11].

(校对责编: 牛欣悦)