



航空动力学报

Journal of Aerospace Power

ISSN 1000-8055, CN 11-2297/V

《航空动力学报》网络首发论文

题目: 基于卷积双向长短期网络的轴承故障尺寸估计
作者: 刘西洋, 陈果, 郝腾飞, 潘文平
DOI: 10.13224/j.cnki.jasp.20210292
收稿日期: 2021-06-09
网络首发日期: 2022-05-30
引用格式: 刘西洋, 陈果, 郝腾飞, 潘文平. 基于卷积双向长短期网络的轴承故障尺寸估计[J/OL]. 航空动力学报. <https://doi.org/10.13224/j.cnki.jasp.20210292>



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。

基于卷积双向长短期网络的轴承故障尺寸估计

刘西洋¹, 陈果^{1,2}, 郝腾飞³, 潘文平¹

(1. 南京航空航天大学 民航学院, 南京 211106;

(2. 南京航空航天大学 通用航空与飞行学院, 南京 213300)

3. 南京工程学院 汽车与轨道交通学院, 南京 211167)

摘要：基于振动监测数据的航空发动机滚动轴承损伤大小识别，对于研究滚动轴承故障演化、故障预测和故障诊断具有重要意义。针对传统模型对先验知识依赖性高、特征提取不充分、故障尺寸训练类别有限等问题，本文提出了一种基于深度学习的滚动轴承损伤尺寸预计方法，能够对训练过程中未出现的中间尺寸进行准确识别。在经典模型的基础上，搭建了一种深度卷积网络与长短期记忆网络组合模型，该模型可对轴承振动信号的多维特征与时序特征进行充分提取，实现轴承故障的智能和高效诊断。最后，利用滚动轴承加速疲劳试验机，进行了多种转速与损伤尺寸下的滚动轴承故障试验，基于试验数据进行了方法的比较，结果表明，该组合网络的在正常和加噪的情况下预测精度分别达到 99.94% 和 98.67%，较单独的深度卷积网络、长短期记忆网络及其他模型精度更高，比较结果充分表明了本文所提方法的优越性。

关键词：滚动轴承；故障诊断；损伤尺寸；深度卷积网络；长短期记忆网络

中图分类号：V263.6 **文献标志码：**A

Rolling Bearing Fault Size Estimation Based on Convolutional Bidirectional Long and Short Term Memory Network

LIU Xiyang¹, CHEN Guo^{1,2}, HAO Tengfei³, PAN Wenping¹

(1. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;

2. College of General Aviation and Flight, Nanjing University of Aeronautics and Astronautics, Nanjing 213300, China;

3. School of Automotive and Rail Transit, Nanjing Institute of Technology, Nanjing 211167, China)

Abstract The damage size identification of aero-engine rolling bearing based on vibration monitoring data is of great significance for the study of rolling bearing fault evolution, prediction and diagnosis. There are inherent restrictions in traditional identification models such as high dependence on prior knowledge, insufficient feature extraction and limited category of training fault sizes, aiming at the limitation, a prediction method of rolling bearing damage size based on deep learning is proposed, which can accurately identify the middle sizes that does not appear in the training process. In this paper, a combined model of deep convolutional long-short-term memory network is developed, which can sufficiently extract the multi-dimensional and time-series characteristics of bearing vibration signal, and realize the intelligent and efficient diagnosis of bearing fault. On the basis of theoretical analysis, the rolling bearing fault tests under various damage sizes and rotational velocities are carried out by using the accelerated fatigue testing machine for rolling bearings, and the traditional and novel methods are compared based on the test data. The results show that the prediction accuracy of the combined network can reach 99.94% and 98.67% respectively under normal and noisy conditions, which is higher than the single deep convolution network, long-short-term memory network and other models. The

收稿日期：2021-06-09

基金项目：国家科技重大专项（J2019-IV-004-0071）

作者简介：刘西洋（1994—），博士生。E-mail: lxycca@nuaa.edu.cn

通信作者：陈果（1972—），男，教授，博士生导师。E-mail: cgzyx@263.net

comparison results amply demonstrate the superiority of the proposed method.

Key words: Rolling element bearing; Fault diagnosis; Damage size; Deep convolution network; Long-short-term memory network

滚动轴承是航空发动机的重要支承部件,对航空发动机的使用安全、寿命和可靠性具有重大影响。滚动轴承疲劳剥落是一种主要的失效形式,其剥落面积的大小直接反应滚动轴承的状态。试验规程规定,在滚道或滚动体上出现面积为 0.5mm^2 的疲劳剥落坑就认为轴承寿命终结。然而,在实际运行过程中无法直接测量出滚动轴承的损伤尺寸大小,且剥落故障具有一定时期的演化过程,通常轴承的损伤大小超过两个滚动体间距时,才定义为轴承失效。因此,基于振动监测数据,在剥落故障演化过程中,损伤大小低于滚动体间距的情况下识别出损伤大小,对于提前诊断轴承故障和预测轴承剩余寿命具有重要意义。

国内外学者对此进行了较为广泛的研究,主要研究工作体现在两方面,1)直接基于振动信号的时域波形特征进行损伤大小识别;2)基于振动监测数据机器学习的状态评估方法。在基于振动信号时域波形的损伤大小识别方面,Luo 等人^[1]基于滚珠经过表面损伤区的运动学机理以及 Hertzian 接触理论,建立了以双脉冲时间间隔 (Dualimpulse time separation, DTS) 和滚动体相互作用关系为函数的损伤尺寸估算模型;Kogan 等人^[2]提出一种滚动体与损伤外圈相互作用的多体非线性动力学模型,在考虑径向载荷、转速与重力的条件下建立了损伤尺寸随冲击时间变化的函数表达式,并研究了表达式对几何参数的敏感性;XU 等人^[3]将奇异值分解与复合平方包络谱结合,实现了微弱故障信号提取与故障识别;Niehaus 等

人^[4]提出一种新颖信息准则(NIC)法对信息频段进行自动选择,该方法可以在高波动及低信噪比的情况下实现轴承损伤识别;

基于机器学习方法的轴承损伤特定尺寸识别相关研究相对较少。在机器学习状态评估方法方面,Wang 等人^[5]提出一种残差和长短期记忆网络的组合网络,将方位特征信息加载到一个长时记忆单元中,并引入遗忘机制来提取时间序列数据的全局特征;Dovedi 等人^[6]将轴承振动信号进行经验模态及可调 Q 因子小波时频二次分解,并提取基于分形维数的特征作为时频子带复杂度量,结合支持向量机对其诊断效果进行了评估;Toma 等人^[7]将离散小波变换与机器学习算法进行集成,从小波分解后的系数中提取特征并通过随机森林和极值梯度 boost 进行训练,结果表明该方法较其他模型具有更高的精确度。

然而,直接基于振动信号的时域波形特征分析需要建立在非常明显的冲击响应特征基础上,受实际噪声干扰严重,分析方法的可靠性和稳健性很低,难于实际应用;基于振动信号机器学习的方法往往只能对滚动轴承的状态做出定性的判断,很少直接应用于获取轴承表面的损伤尺寸预计,且传统的机器学习方法往往受学习样本和特征提取方法的优劣影响非常大。

深度学习方法在特征自适应提取方面有着巨大的优势,因而被广泛关注与讨论,成为近年来研究的热点方向^[8]。近年来,随着深度学习的火热而兴起的卷积神经网络 (Convolutional neural network, CNN) 方法,

以其在大规模数据特征提取方面的独特优势而受到广泛关注。CNN 的核心内容是通过卷积与池化运算,使网络能够学到模式的空间层次结构且具有平移不变性^[9]。在 CNN 基础上发展而来的深度 CNN 网络,相比浅层网络具有更强大的学习能力,因而广泛应用于模式识别、故障诊断等领域。长短期记忆网络(Long and short term memory network, LSTM)是在循环神经网络的基础上发展而来的,其独有的“门结构”在一定程度上解决了梯度爆炸或消失问题,能保留长期有效信息,对于时序数据处理具有独特的优势^[10]。

考虑到所使用的模型中,利用 CNN 对原始时序数据进行自适应特征提取只考虑了数据在多维度的特性,未考虑数据的时间序列特性,因此会造成原始数据序列特征信息的丢失;采用 LSTM 神经网络虽然可以对原始序列数据建模并进行序列特征提取,但忽略了时序数据在多维度的特征。因此本文将 CNN 与 LSTM 网络进行组合,首先通过 CNN 自适应地从原始信号中提取故障特征,再由 LSTM 网络对特征进行学习以建立时序回归模型,实现滚动轴承故障特征的自动提取与损伤大小估计。最后通过实际试验数据进行方法验证。

1 滚动轴承损伤大小识别的基本思想

滚动轴承损伤大小识别的本质是基于数据的回归分析,是一个从多维实数空间到一维实数空间的映射。多维实数向量代表滚动轴承监测振动信号的时间序列数据,一维实数及代表输出的损伤大小。因此,滚动轴承损伤大小识别的关键在于两方面,即:1) 如何从振动监测信号中提取出最优质的特

征;2) 如何构建更好的回归预测模型用以对损伤大小的识别和预计。研究流程如图 1 所示。

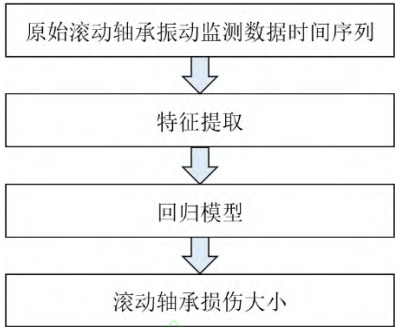


图 1 滚动轴承损伤大小识别流程

Fig. 1 Flowchart of identification of rolling bearing fault size

2 基于卷积长短期记忆网络的轴承损伤估计方法

为同时考虑到轴承故障信号的多维度特性与时序特性,本文提出了一种基于长短期记忆卷积神经网络的滚动轴承损伤尺寸预测模型。该模型由深度卷积神经网络与双向 LSTM(Bidirectional long and short term memory network, BiLSTM)网络组成,能够对输入的原始振动信号进行深层特征和时间动态信息进行学习。深度卷积网络能够自适应地从原始信号中捕获轴承故障的空间特征并减少冗余数据,BiLSTM 网络则负责提取数据的顺序时间特征,将 CNN 提取的空间状态进行连接,并实现滚动轴承损伤尺寸的预测。

2.1 特征提取

卷积神经网络是一种多层感知器神经网络,其采用权值共享的卷积,通过增加网络深度,可以挖掘更丰富的数据信息^[11]。作为深度学习中最常用的方法之一,与传统神经网络相比,深度卷积神经网络最大的优势在于其可以自动从数据中学习到的更为抽象

的特征,并对特征进行计算以用于分类或回归问题。

2.1.1 激活函数

在 CNN 网络传播的过程中需要更新的参数包括输入值 x_i , 权重值 ω_i , 偏置 b , 以及卷积核等。激活函数包括 ReLU、sigmoid、tanh 等,用来为神经网络引入非线性。卷积运算过程如式(1)所示。

$$y_j = f\left(\sum_{i=1}^n \omega_{ij}x_i + b_j\right) \quad (1)$$

其中 x_i 表示输入信号, y_j 表示输出信号, ω_{ij} 表示神经元 j 与输入信号 x_i 所连接的权值, b_j 表示偏置。 $f(\bullet)$ 为激活函数, 本文所用激活函数为 ReLU, 其公式为:

$$f(y_j) = \max(0, y_j) \quad (2)$$

神经网络单元内部传递结构如图 2 所示。

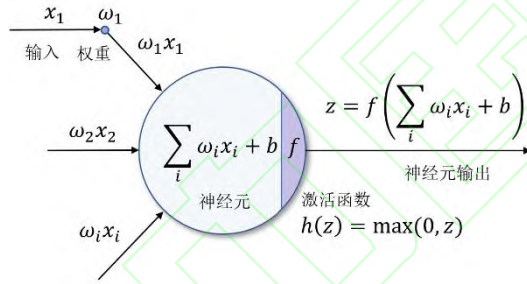


图 2 神经网络单元结构

Fig. 2 Unit structure of neural network

2.1.2 网络结构

本文搭建深度卷积网络模型由输入层、卷积层(C1-C3)与池化层(P1-P3)、dropout 层、BN 层、全连接层(F1)与回归层组成。网络各层具体介绍如下。

1) 卷积层与池化层。卷积层用于对输入图像进行特征提取, 卷积核数量分别为 16、32、32。池化层共有 3 层, 池化方式采

用平均池化, 负责对输入图像特征进行压缩与降维。二者交错堆叠, 以完成数据降维与特征提取。本文在输入层后连接一层 2×2 的平均池化层, 在减少计算量的同时尽可能保留更多的图像信息。模型的激活函数采用 ReLU, 相对于其他激活函数, ReLU 具有稀疏特性且梯度稳定, 在进行梯度下降和误差反向传播时有着更高的效率。

2) Dropout 层。在全连接层前添加 dropout 层, 以 0.2 的概率随机置零神经元, 以期望提高模型的泛化能力。Dropout 为 Srivastava 等人提出的一种正则化方法, 其主要思想为增加对特征的随机选择, 减少模型的过拟合^[12]。相比其他正则化方法, dropout 方法具有计算方便、不依赖修改代价函数、对模型适应力强等特点, 能够显著提升过拟合问题突出的深度网络的性能。

3) BN (Batch Normalization) 层。在输入层后和 dropout 层前各添加一层 BN 层, 目的在于进一步加快网络训练速度, 防止梯度爆炸或消失以及减小过拟合。

4) 全连接层。全连接层位于网络的核心特征提取架构层与回归层之间, 将最后一层输出的局部特征信息进行综合并输入至回归输出层, 其大小设置为 1。

5) 回归层。与常用的分类网络不同的是, 本文所采用的回归网络在全连接层后不设置激活函数, 直接将尺寸判断结果以回归函数输出, 最大程度保留特征提取结果。

所搭建网络结构如图 3 所示。

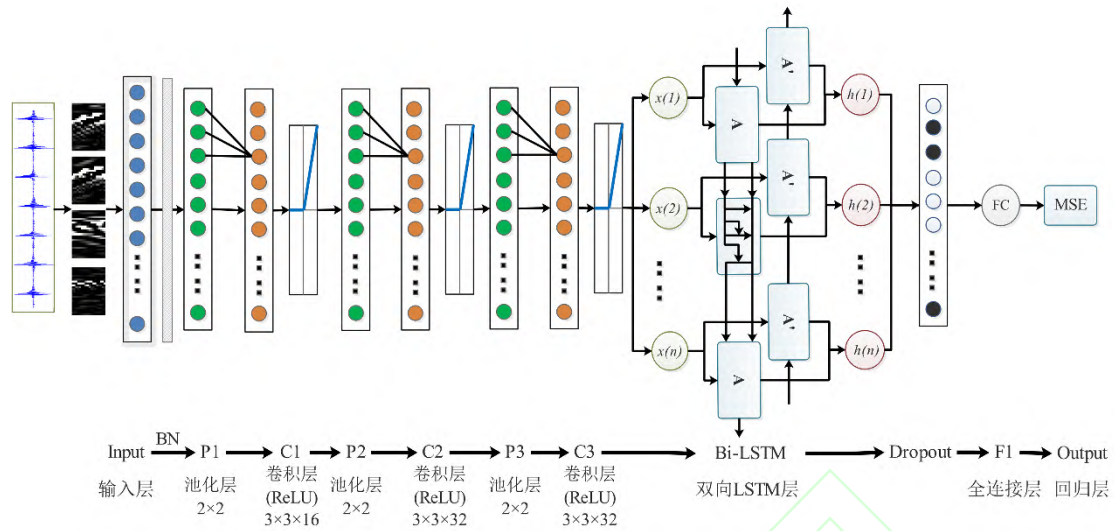


图 3 DCNN+BiLSTM 网络结构示意图

Fig. 3 Structural diagram of DCNN+BiLSTM network

2.1.3 损失函数与优化器

神经网络的训练需要选择适合的损失函数与优化器，通过迭代更新寻找最优参数。由于本文使用回归网络，网络输出为具体的尺寸数值，因此采用均方根误差(Root mean square error, RMSE)作为网络的损失函数，以直观描述预测值与真实值的差距。优化器则用以对损失函数进行优化，常用的优化方法有梯度下降法，动量优化法，自适应学习率优化法等。本文比较了在学习率为 0.001，迭代轮次为 30 的情况下不同优化器的表现，如表 1 所示。

表 1 不同优化器的训练时长与损失对比

Table 1 Training duration and verification loss of different optimizers

优化器	批量大小	最终验证损失	训练耗时 (s)
Sgdm	32	0.0094	426
	64	0.0022	275
	128	0.0061	192
Adam	32	0.0261	583
	64	0.0193	404
	128	0.0270	239
RMSProp	32	0.0221	624
	64	0.0272	621

	128	0.0134	507
--	-----	--------	-----

通过对比可以发现，Sgdm 优化器在滚动轴承故障验证集上的表现最好，最终验证损失最低。因此本文选用 Sgdm 优化器，批量大小设置为 64。

2.1.4 CNN 特征提取流程

将不同损伤尺寸的振动数据，转换成矩阵图的形式，输入 25×18 的神经元中，这些神经元负责对矩阵图的像素亮度进行编码；卷积层使用 3×3 的局部感受野对图像进行原始特征捕捉，池化层负责特征的降维和进一步提取，形成一系列凝缩的特征映射。随着卷积层和池化层的交叠使用，能够组合形成更为抽象的特征，最终形成对不同图像的特征描述。另外，卷积与池化的局部视野感受特点，使得损伤特征提取不受其所在图像位置的影响，即 CNN 网络能够在尺度空间不变性的基础上完成不同损伤尺寸的特征提取。

使用卷积神经网络提取轴承损伤特征的流程如图 4 所示。其中不同灰度的像素表明了网络所学习权重的大小，像素响应的增加表明神经网络学习到了与空间结构相关

的更深层内容。

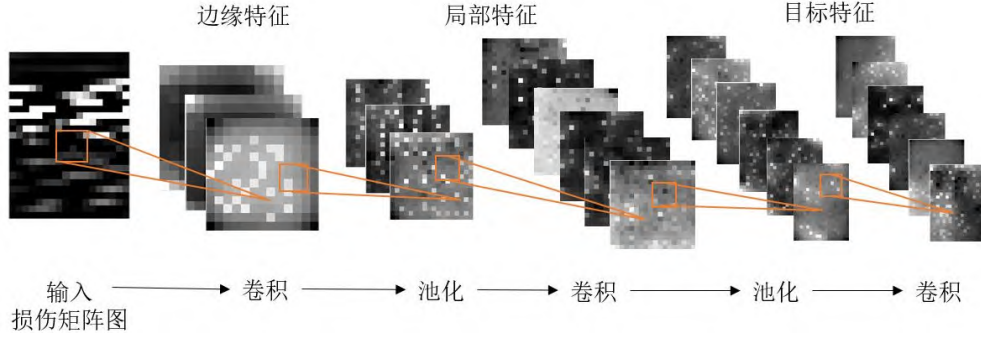


图 4 卷积神经网络特征提取流程

Fig. 4 Feature extraction illustration of CNN network

2.2 故障尺寸估计

卷积神经网络通过卷积块在图像上的平移来提取特征，能够描述数据的某一空间状态，其本身不具有记忆功能，无法考虑数据与时间的相关性，即对时间不敏感。滚动轴承振动信号作为一种时间序列数据，在进行特征提取与尺寸估计时应考虑其时间关联特性。长短期记忆网络是一种递归神经网络^[13]，其在神经网络的基础上添加了时间序列与定向循环的概念，能够对长期顺序依存关系进行学习。与传统循环神经网络相比，LSTM 网络能够保存较早的信息且实现跨时间步的传送，因而能够解决传统网络中存在的梯度不稳定问题^[14]。

LSTM 单元内部的计算方式如下：

$$\begin{aligned}
 f_t &= \sigma(W_f \cdot (h_{t-1}, x_k^{(t)}) + b_f) \\
 i_t &= \sigma(W_i \cdot (h_{t-1}, x_k^{(t)}) + b_i) \\
 \tilde{c}_t &= \tanh(W_c \cdot (h_{t-1}, x_k^{(t)}) + b_c) \\
 c_t &= f_t c_{t-1} + i_t \tilde{c}_t \\
 o_t &= \sigma(W_o \cdot (h_{t-1}, x_k^{(t)}) \cdot x_k^{(t)} + u^{(o)} \cdot h_{t-1}) \\
 h_t &= o_t \tanh(c_t)
 \end{aligned} \quad (3)$$

式中 W 和 b 分别为模型学习得到的权重矩阵和偏置向量， $x_k^{(t)}$ 为 t 时刻输入向量， $\sigma(\bullet)$ 和 $\tanh(\bullet)$ 分别为 sigmoid 激活函数与 tanh 激活函数。LSTM 网络通过遗忘门 f_t ，

输入门 i_t 和输出门 o_t 组成的 LSTM 单元，以及 $t-1$ 时刻的隐层状态 h_{t-1} 和细胞状态 c_{t-1} 进行更新，得到当前的细胞状态 c_t ，进而更新细胞中的当前隐层状态 h_t ，并依次传递给下一单元。每个 LSTM 单元结构如图 5 所示：

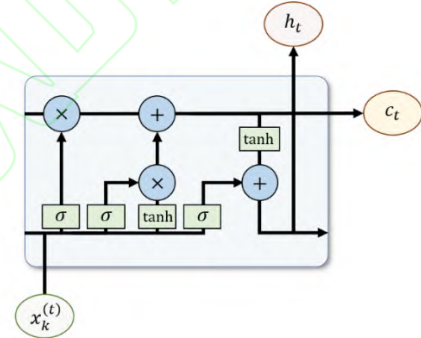


图 5 LSTM 单元结构

Fig. 5 Unit structure of LSTM net

本文通过将每个图像输入至深度卷积神经网络模型进行特征提取后，再将每个输出作为单位时间步长传递给 LSTM 网络进行回归预测。为实现这一过程，本文将整个 CNN 网络封装在一个时间分布层中以实现多次运用，并依次向后续 LSTM 网络提供一系列所提取的图像特征。

在对滚动轴承损伤某个未知尺寸进行估计时，需要同时结合该尺寸前后的已知尺寸。同样，在同一个尺寸的振动序列中不同冲击间也存在相应关联，在预测时需要做到基于上下文同步判断。单向 LSTM 网络仅能

根据之前时刻序列信息对后面信息进行预测，反之则无法进行。

为了同步考虑到轴承振动信号的上下文信息，更好地保证 LSTM 网络的特征提取性能，本文搭建一种双向 LSTM 网络。BiLSTM 网络由两个方向相反的 LSTM 层组成，前向传播层与后向传播层分别从序列的首段和尾端开始，逐层递进传播，二者共同连接输出层，实现权值共享，最终对得到的两个结果进行同步处理。该网络可同时获取过去和将来的信息，进一步缓解遗忘问题，提高预测精度。在完成特征提取后，利用网络最后一层的回归函数作为输出层，用基于实数值的误差作为优化目标函数，进行 DCNN (Deep Convolution Neural Network) +BiLSTM 回归网络的训练。

完成模型的训练后，选择未参与训练的中间未知尺寸的振动监测信号的时间序列数据作为测试集输入至该模型，得到估计结果。将该估计结果与真实标签进行比较，得到二者的误差值，作为衡量学习优劣的标准。性能评估标准主要有三点：1) 在可接受误差界限内的预测值的百分比；2) 估计尺寸大小和实际尺寸大小的均方根误差；3) 估计值与实际值的相对误差。

3 滚动轴承不同损伤的振动试验

3.1 试验设备

通过上述分析表明，为了实现基于深度学习的滚动轴承损伤大小识别，需要利用大量的样本对 DCNN+BiLSTM 回归网络进行训练和测试。因此，本文利用试验的方法来得到大量不同滚动轴承损伤大小的真实振动监测数据。

本试验用到的试验设备包括 AB-LT1A 型轴承寿命强化试验机、AI002 加速度传感

器、JM5937 动态信号测试分析系统等。轴承寿命强化试验机主要用于滚动轴承疲劳寿命强化（快速）试验，其结构如图 6 所示。该试验机在试验中需要安装 4 个轴承，本次试验中，每次试验安装 1 个故障轴承和 3 个正常轴承。

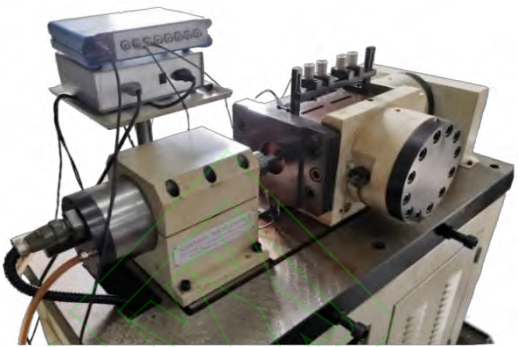


图 6 轴承寿命强化试验机

Fig. 6 Bearing life strengthening testing machine

试验共使用了 9 个 HRB 6206 深沟球轴承，该型号轴承的主要参数如表 2 所示。采用电火花线切割技术，人为在其外圈表面分别加工了 9 种不同宽度的凹槽来模拟不同故障尺寸，凹槽宽度依次为 0.8mm、1.0mm、1.2mm、1.4mm、1.6mm、1.8mm、2.0mm、2.2mm 和 2.4mm，故障加工细节如图 7 所示。试验时依次装入 9 个轴承，利用轴承座上的加速度传感器采集振动信号，转速设置为 2400 r/min，采样频率设置为 128kHz。



图 7 轴承故障加工细节

Fig. 7 Processing details of bearing fault

表 2 HRB 6206 深沟球轴承的主要参数
Table 2 Main parameters of HRB 6206 deep groove ball bearing

内径	外径	厚度	滚珠直径	节径	滚珠数	接触角
30 mm	62 mm	16 mm	9.5 mm	46 mm	9	0°

3.2 试验数据

3.2.1 数据采集

本试验共采集了9种尺寸的滚动轴承故障数据，每种尺寸下约 50 个样本，每个样本包含 131072 个采样点，时间长度为 1s。9 种损伤尺寸下的轴承振动时域波形如图所 8 示。从图中可以看出，仅仅依靠振动冲击的幅值大小无法准确判断损伤的尺寸大小。而需要从信号冲击的时域波形的时间序列数据中提取特征信息。这也是本文利用深度学习模型进行滚动轴承损伤大小识别的根本出发点。

数据采集具体方案如表 3 所示。

表 3 数据采集的具体方案

Table 3 Specific scheme of data acquisition	
项目	对应参数
采样频率 (kHz)	128
转速(r/min)	2400
训练尺寸 (mm)	0.8、1.0、1.2、1.6、2.0、2.2、2.4
预测中间尺寸(mm)	1.4、1.8
采样时间(s)	1
数据点数	131072
样本数	每种故障 50 个

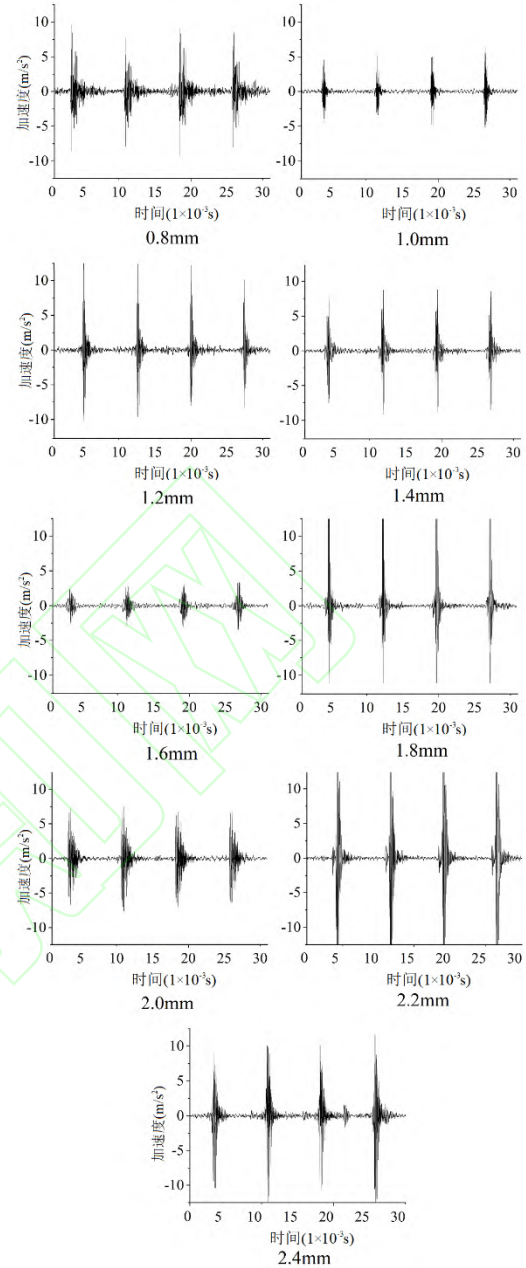


图 8 滚动轴承 9 种损伤尺寸时域波形图

Fig. 8 Time domain waveform of 9 fault sizes of rolling bearings

3.2.2 故障轴承的双冲击特征

为了建立切实有效的深度学习模型，需要从机理上认识滚动轴承表面故障所带来的冲击信号规律和特征，滚动轴承发生疲劳剥落故障后，会使得滚珠在经过损伤区域边缘时产生振动响应。Epps 与 McCallion^[15]最先对此进行了研究，指出滚珠在进入剥落区前边缘时，振动信号将呈现出阶跃响应特

性，且以相对较低的频率成分为主；撞击剥落区后边缘时，振动信号则呈现出脉冲响应特性，产生高频响应，并随着滚珠的离开产生低频衰减，其原理如图 9 所示。Randall^[16]和 Sawalhi^[17]也通过研究表明，滚珠进入和离开剥落区的两个冲击峰值间隔与故障尺寸成正比，可作为故障大小的度量。

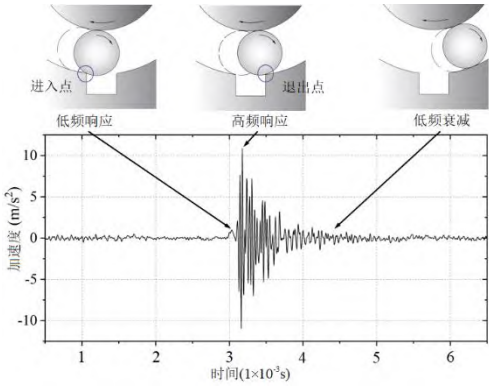


图 9 滚动体进入至离开剥落区所产生的振动响应
Fig. 9 Vibration response of rolling element entering and leaving fault zone

理论上，轴承滚动体在经过损伤区域时存在清晰的双冲击现象(Double impulses phenomenon)，且冲击时间间隔随着轴承损伤尺寸的增大而增大。但在航空发动机实际工作过程中，轴承振动信号往往来自多个振源且受背景噪声等因素影响，使得所采集信号的冲击特征不够明显，人为提取方法及一般信号处理方法难以对双冲击现象进行有效分离。针对该问题，本文采用深度学习方法，从时域样本中对故障信号这一特征进行提取，以期望自动识别轴承损伤尺寸规律。

3.2.3 数据预处理

与一维卷积网络相比，二维卷积网络在降低神经网络的训练难度的同时，能够提高网络模型的诊断准确率。相关研究^[18,19]指出，尽管目前深度学习在计算机领域发展迅速且应用广泛，但在故障诊断领域，数据大

多为一维时间序列，对预测模型的构建来说具有难度。原因主要包括一维网络训练难度高且精度低，现存预训练网络不足等等，因而不如二维卷积网络具有普适性。因此本文将轴承振动数据编码为二维图片格式，采用 2D 卷积神经网络进行训练。

为了提高样本的质量，更好地提取故障特征，根据滚动轴承故障的周期冲击原理，本文将每个振动监测信号的时间序列进行了预处理。具体步骤为：1) 以信号中每次滚珠冲击故障点产生的峰值，即每个冲击周期的最高峰为原点，前后分别截取 150 和 300 个数据点，以此 450 个数据点作为一个学习样本；2) 遍历每个时间序列的所有冲击周期，即可得到多个学习样本；3) 最终采用该方式对多个包含 9 种不同损伤尺寸的振动信号的时间序列数据进行处理，共计得到不同损伤尺寸下的学习样本各 2600 个；4) 通过连续交错采样对每个学习样本进行排序，生成数据矩阵，形成 25×18 的图像。该方法在信号分段后将未经其他处理的原时域信号作为输入，避免噪声干扰的同时最大程度地保留信号中的信息，并借由数据驱动模型进行自动特征提取。方法流程如图 10 所示。

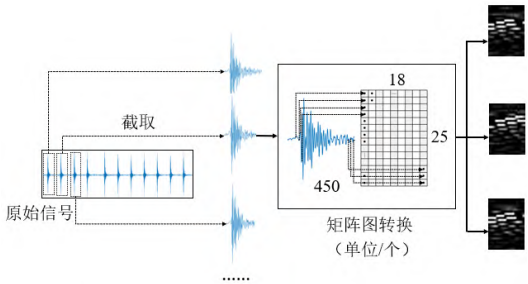


图 10 学习样本预处理流程

Fig. 10 Preprocessing of learning samples

根据上述方法，将 7 种损伤尺寸的数据作为训练集，2 种中间尺寸的数据不参与训练，作为预测集，实现滚动轴承故障大小的

回归预测。同时评估预测结果的精确度。

4 基于试验数据的方法比较

4.1 机器学习模型估计结果

考虑到传统的人为特征提取方法可能导致原始信号中包含滚动轴承故障尺寸的信息丢失,下面将分别使用支持向量回归模型、长短期记忆神经网络模型、深度卷积神经网络模型、深度卷积网络与长短期记忆网络组合模型,分别对轴承原始振动信号进行故障特征的自动提取,构建回归预测模型,并进一步进行中间尺寸的预测。截取后的每450个点构成一个训练样本,其中7种尺寸用来训练模型,2种中间尺寸作为预测对象,最终组成训练集18200个,测试集5200个。

4.1.1 支持向量回归预测结果

搭建支持向量回归预测模型,选取径向基核函数,惩罚系数 $C=1$, γ 设置为模型自动选取,得到预测结果如图11所示。

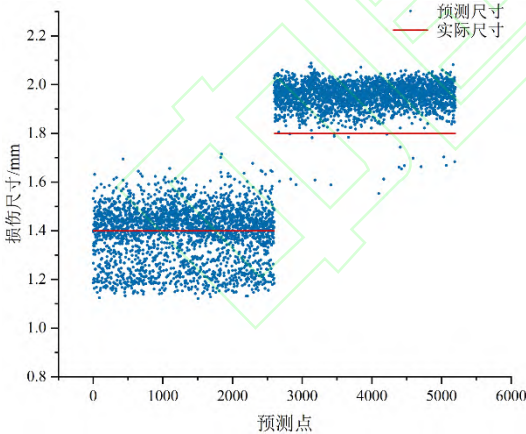


图 11 支持向量回归预测结果

Fig. 11 Predicted results of Support Vector Regression

4.1.2 双向长短期记忆网络预测结果

增加 LSTM 网络的层数有利于提高网络深度,因此本文搭建一双向 LSTM 结构,以实现更好的特征提取。隐藏节点个数设置为200,最终得到预测结果如图12所示。

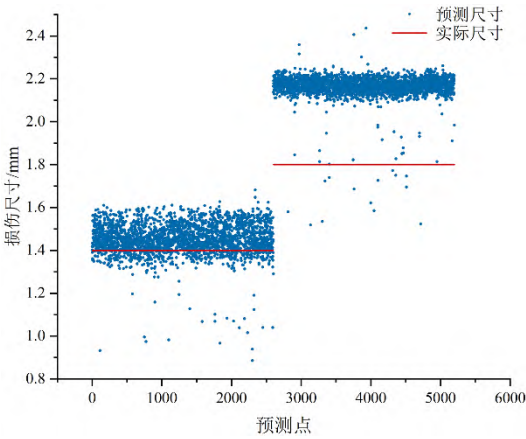


图 12 双向长短期记忆网络(BiLSTM)预测结果

Fig. 12 Predicted results of bidirectional long and short term memory(BiLSTM) network

从图11和图12可以看出,与SVR网络相比,LSTM网络预测结果的分布散度更低,但是对个别尺寸预测偏差较大,因而导致整体预测率偏低。

4.1.3 深度卷积神经网络预测结果

按照前述理论,搭建深层卷积神经网络模型。其中优化器选择SGDM,即在SGD的基础上引入了一阶动量,训练轮数设置为30。将初始学习率设置为0.001,并在20轮训练后逐渐降低10%。训练过程中设置验证数据,用来在训练过程中按固定时间间隔计算准确度,但不参与网络权重的更新。

DCNN模型预测结果如图13所示。

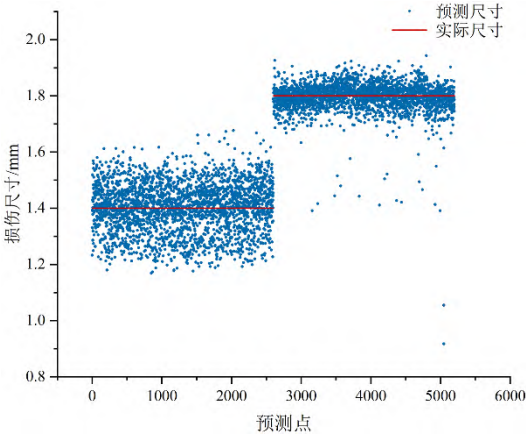


图 13 深度卷积神经网络(DCNN)预测结果

Fig. 13 Predicted results of deep convolution neural network (DCNN)

从深度卷积神经网络预测结果可以看出，该网络与之前两种网络相比，预测精度有了很大的提高，但是相比 BiLSTM 网络来说，预测结果较为分散，因此将二者进行组合，采用 DCNN+BiLSTM 网络进行模型训练与预测。

4.1.4 长短期记忆卷积神经网络

将上述深度卷积神经网络与长短期记忆网络结合，组成 DCNN+BiLSTM 网络，其中 DCNN 部分主要由 3 个卷积层与池化层组成，在最后一层激活函数后连接双向 LSTM 层，网络其他参数保持不变。最终获得训练结果如图 14 所示。

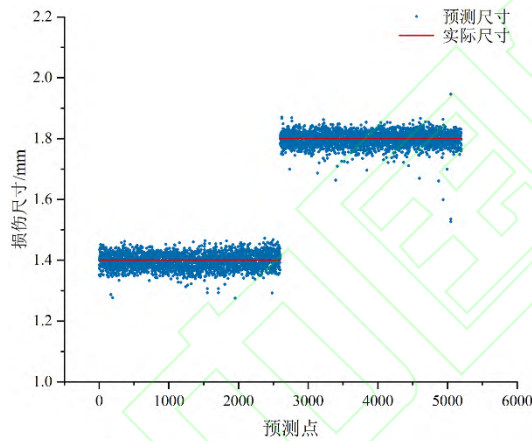


图 14 DCNN+BiLSTM 组合网络预测结果

Fig. 14 Predicted results of DCNN+BiLSTM combined network

从 DCNN+BiLSTM 组合模型的预测结果可以看出，该预测数据的预测精度和集中度都高于前述模型，异常点数量较少，证明该组合能够更有效地对轴承故障特征进行提取，并对未参与训练的中间数据进行准确预测。

4.2 试验结果对比

4.2.1 衡量指标

为了进一步对各模型预测结果进行具体分析，对预测数据提取 3 种指标：精确度、均方根误差 RMSE 以及相对误差 δ 。由于本文所列举损伤尺寸间隔大小为 0.2mm，设待识别的尺寸为 a ，则识别结果在 $a \pm 0.2\text{mm}$ 范围内均认为正确，因此定义精确度阈值为尺寸 a 的变化范围，即 $\pm 0.2\text{mm}$ ，超过了范围则视为识别错误，本文用该精确度来衡量判定正确的结果所占比例，同时，定义 0.2mm 的识别误差容限也具有实际工程意义，实际轴承失效的损伤大小将达到数毫米。RMSE 用于衡量预测值与实际值之间的偏差，作为线性回归的损失函数，也常用于机器学习模型预测结果的评判标准。与均方误差相比，RMSE 能够在同一数量级更直观地描述具体误差值，RMSE 值越低，证明预测误差越小；相对误差 δ 则用来衡量预测误差占真实值的百分比，对于轴承故障尺寸的估计问题，与 RMSE 值相结合，能够更加直观可信地反映所预测尺寸的精度。

RMSE 及 δ 的计算方式如下：

$$R(y_i, \hat{y}_i) = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (4)$$

$$\delta = \frac{|y_i - \hat{y}_i|}{y_i} \quad (5)$$

通过计算，得到转速 2400r/min，预测尺寸为 1.4mm 和 1.8mm 的情况下，几种基于机器学习模型的预测结果对比如表 4 所示。

表 4 1.4mm 与 1.8mm 损伤估计结果对比
($n=2400\text{r/min}$)

Table 4 Comparison of estimation results of 1.4mm and 1.8mm fault sizes ($n=2400\text{r/min}$)

模型	精确度	RMSE	相对误差
----	-----	------	------

SVR	86.42%	0.3731	0.0741
DCNN	98.94%	0.0772	0.1547
BiLSTM	49.77%	0.3392	0.0339
DCNN+BiLSTM	99.94%	0.0253	0.0125

从 3 项衡量指标可以看出,在几种单独机器学习特征提取方法中,DCNN 模型在误差 0.2mm 内预测值所占百分比更高,且均方根误差与相对误差均较低。与单独模型相比,DCNN+BiLSTM 模型的指标衡量结果均优于其他模型,证明其在特征提取方面的优越性。

为充分验证 DCNN+BiLSTM 模型的损伤尺寸估计性能,本文将同转速下的另外两个尺寸(1.2mm、2.0mm)作为预测集,其他尺寸作为训练集,得到预测结果如表 5 所示。

表 5 1.2mm 与 2.0mm 估计结果对比 ($n=2400\text{r/min}$)

Table 5 Comparison of estimation results of 1.2mm and 2.0mm fault sizes ($n=2400\text{r/min}$)

模型	精确度	RMSE	相对误差
SVR	79.77%	0.1575	0.0765
DCNN	98.88%	0.0605	0.0281
BiLSTM	69.12%	0.2603	0.1324
DCNN+BiLSTM	99.48%	0.0530	0.0252

从表 4 可以看出,与预测 1.4 mm 和 1.8 mm 的情况类似,在预测尺寸为 1.2 mm 和 2.0 mm 时,DCNN+BiLSTM 模型也具有最高的精确度和最低的误差值。

4.2.2 其他转速情况

考虑到不同转速下轴承振动信号状态会发生相应变化,为验证本文所述模型在其他转速下的泛化能力,采集了滚动轴承在 1200r/min 下的振动信号,分别对两组不同损伤尺寸进行预测,得到预测结果如表 6、表 7 所示。

表 6 1.4mm 与 1.8mm 预测估计对比 ($n=1200\text{r/min}$)

Table 6 Comparison of estimation results of 1.4mm

and 1.8mm fault sizes ($n=1200\text{r/min}$)

模型	精确度	RMSE	相对误差
SVR	87.37%	0.1516	0.0871
DCNN	97.92%	0.0645	0.0400
BiLSTM	92.08%	0.1215	0.0637
DCNN+BiLSTM	99.44%	0.0435	0.0266

表 7 1.2mm 与 2.0mm 预测估计对比 ($n=1200\text{r/min}$)

Table 7 Comparison of estimation results of 1.2mm

and 2.0mm fault sizes ($n=1200\text{r/min}$)

模型	精确度	RMSE	相对误差
SVR	88.87%	0.1371	0.0728
DCNN	96.75%	0.0752	0.0297
BiLSTM	83.25%	0.1859	0.0918
DCNN+BiLSTM	99.77%	0.0495	0.0246

从 1200r/min 转速下的两组不同尺寸预测结果可以看出,与 2400r/min 转速下的情况类似,DCNN+BiLSTM 组合模型在损伤尺寸大小估计方面的准确度要高于两种单独的模型。

4.2.3 考虑噪声的情况

本实验数据为轴承座测点直接采集所得,然而实际航空发动机的传感器测点通常位于涡轮机匣附近,信号传递路径复杂且包含多种噪声耦合,故障特征衰减严重,对于传统故障诊断方法来说,往往难以达到理想的特征提取效果。为考察本文所述模型在含噪情况下对于损伤尺寸估计的精度,为上述采集信号添加均值为 0,方差为 1 的高斯白噪声,来模拟实际工况下轴承的振动情况。添加噪声前后的滚动轴承振动时域波形图与矩阵图如图 15 所示。

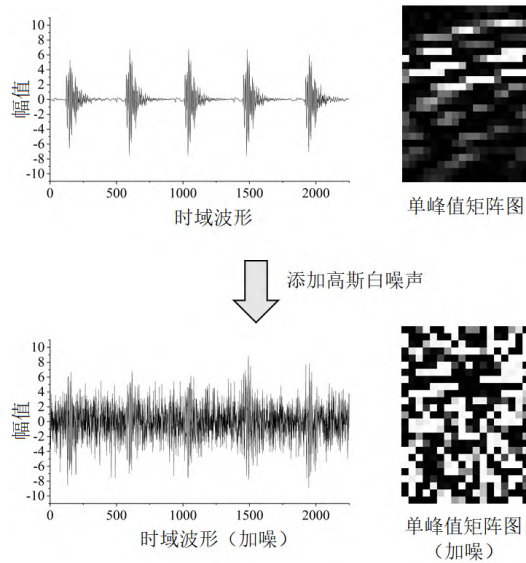


图 15 添加高斯白噪声前后的信号时域波形与矩阵图

Fig. 15 Time domain waveform and matrix diagram of signal before and after adding white Gaussian noise

分别用四种机器学习模型对加噪信号进行损伤尺寸估计，得到估计结果如表 8 所示。

表 8 加噪声情况下 1.4mm 与 1.8mm 估计结果对比
Table 8 Comparison of estimation results of 1.4mm and 1.8mm fault sizes by adding noise (2400r/min)

模型	精确度	RMSE	相对误差
SVR	50.15%	0.1772	0.1058
DCNN	77.50%	0.1640	0.0824
BiLSTM	50.41%	0.2010	0.1265
DCNN+BiLSTM	98.67%	0.0846	0.0414

从表 7 中可以看出，在模拟真实工况的条件下，单独机器学习模型的估计精度有明显的下降，且误差值显著提升。与信号不含噪的情况相比，在含噪状态下，三种单独模型的平均精确度下降了 19.02%；相对而言，DCNN+BiLSTM 组合模型的精度下降率小很多，为 0.81%，可以看出该模型受噪声影响较小。

另外，在含噪与不含噪的情况下，本文所搭建模型的预测准确率与误差波动范围均很小，且预测精度始终远高于其他三种模型，证明了该模型在滚动轴承损伤尺寸估计

的实际应用方面具有显著优势。

5 结论

1) 在单独的机器学习模型中，DCNN 网络较 SVR 和 BiLSTM 网络具有更高的精确度，以及更低的均方根误差值及相对误差值，对于轴承损伤尺寸估计结果更好；

2) 在 DCNN 网络和 BiLSTM 网络的基础上搭建一种卷积双向长短期记忆网络，通过实验证明了该网络在轴承损伤尺寸估计方面，较单独网络相比具有更高的估计精度；

3) 随着转速、估计尺寸和噪声的变化，单独模型的估计精度和误差值均出现一定范围的波动，而 DCNN+BiLSTM 组合模型的精确度受轴承转速、损伤尺寸、噪声等因素影响小且准确度维持在较高范围。证明了该模型在滚动轴承损伤尺寸预测方面具有明显优势。

本文方法可为轴承故障演化实时监测、状态评估与健康管理提供重要技术。

参考文献：

- [1] LUO Maolin, Guo Yu, Wu Xing, et al. An analytical model for estimating spalled zone size of rolling element bearing based on dual-impulse time separation[J]. Journal of sound and vibration, 2019, 453: 87-102.
- [2] Kogan G, Bortman J, Klein R. A new model for spall-rolling-element interaction[J]. Nonlinear dynamics, 2017,87(1):219-236.
- [3] XU Lang, Chatterton Steven, Pennacchi Paolo. Rolling element bearing diagnosis based on singular value decomposition and composite squared envelope spectrum[J]. Mechanical systems and signal processing, 2021, 148: 1-25.
- [4] Niehaus W N, Schmidt Stephan, Heyns P S. NIC Methodology: A probabilistic methodology for improved informative frequency band

- identification by utilizing the available healthy historical data under time-varying operating conditions[J]. *Journal of sound and vibration*, 2020, 488: 1-20.
- [5] WANG Youming, CHENG Lin. A combination of residual and long-short-term memory networks for bearing fault diagnosis based on time-series model analysis[J]. *Measurement science and technology*, 2021, 32(1): 1-14.
- [6] Dovedi T, Upadhyay R. Diagnosis of ball bearing faults using double decomposition technique[J]. *International journal of acoustics and vibration*, 2020, 25(3):327-340.
- [7] Toma R N, Kim J M. Bearing Fault classification of induction motors using discrete wavelet transform and ensemble machine learning algorithms[J]. *Applied sciences-basel*, 2020, 10(15): 5251.
- [8] REN Lei, SUN Yaqiang, CUI Jin, et al. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks[J]. *Journal of manufacturing systems*, 2018, 48: 71-77.
- [9] Khan M A, Kim Y H, Choo J. Intelligent fault detection using raw vibration signals via dilated convolutional neural networks[J]. *Journal of supercomputing*, 2020, 76(10): 8086-8100.
- [10] Aljbali S, Roy K. Anomaly Detection Using Bidirectional LSTM[C]// *Intelligent systems and applications*. London: Springer, 2020: 612-619.
- [11] Plakia S, Boutalis Y S. Fault detection and identification of rolling element bearings with Attentive Dense CNN[J]. *Neurocomputing*, 2020,405: 208-217.
- [12] Srivastava Nitish, Hinton Geoffrey, Krizhevsky Alex, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. *Journal of machine learning research*. 2014, 15(1):1929-1958.
- [13] Chemali Ephrem, Kollmeyer Phillip J, Preindl Matthias, et al. Long short-term memory-networks for accurate state of charge estimation of li-ion batteries[J]. *IEEE transactions on industrial electronics*, 2017, 65(8): 6730-6739.
- [14] KONG Weicong, DONG Zhaoyang, JIA Youwei, et al. Short-Term residential load forecasting based on LSTM recurrent neural network[J]. *IEEE transactions on smart grid*, 2017, 10(1): 841-851.
- [15] Epps I K, McCallion H. An investigation into the characteristics of vibration excited by discrete faults in rolling element bearings[C]// *Annual conference of the vibration association of New Zealand*, Christchurch, New Zealand, 1993.
- [16] Randall R B. *Vibration-based condition monitoring: industrial, aerospace and automotive applications*[M]. Australia: John Wiley & Sons, 2011.
- [17] Sawalhi N, Randall R B. Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size[J]. *Mechanical systems and signal processing*, 2011, 25(3):846-870.
- [18] Jastrzebska A. Lagged encoding for image-based time series classification using convolutional neural [J], *Statistical analysis and data mining*, 2020, 13(3): 245-260.
- [19] 张向阳, 陈果, 郝腾飞, 等. 基于机匣信号的滚动轴承故障卷积神经网络诊断方法[J]. *航空动力学报*, 2019, 34(12): 2729-2737.