

基于深度卷积测量网络的滚动轴承压缩域故障特征提取方法

林慧斌, 王洪畅, 习慈羊

(华南理工大学机械与汽车工程学院, 广东 广州 510641)

摘要: 压缩感知可有效降低机械状态监测信号的数据存储和传输压力, 而现有压缩感知方法在故障诊断的应用中存在压缩效率低下、信号重构过程缓慢等问题。本文利用自编码网络与压缩感知的对应关系, 提出了一种基于深度卷积测量网络的滚动轴承压缩域故障特征提取方法。针对无噪声的故障信号样本难以获取的问题, 提出一种利用故障机理构建数据集的方法, 利用该仿真数据集训练得到的模型适用于不同工况下的实测轴承信号。构造网络层数由所需要的信号压缩率确定、隐含层与原信号的频率呈对应关系的深度卷积去噪自编码网络。截取训练完备的编码器网络(即深度卷积测量网络)代替传统的观测矩阵对滚动轴承振动信号进行压缩测量, 实现压缩域的故障特征提取。仿真分析验证了所提数据集构造方法及压缩域特征提取方法的有效性。滚动轴承实验信号分析进一步验证了采用所提方法训练得到的深度卷积测量网络具有很好的泛化性, 且能够在压缩率远低于传统压缩感知方法的情况下有效地提取轴承故障特征成分并进行故障诊断。

关键词: 故障诊断; 滚动轴承; 故障特征提取; 压缩感知; 深度卷积测量网络

中图分类号: TH165⁺.3; TH133.33 **文献标志码:** A **文章编号:** 1004-4523(2024)03-0485-12

DOI: 10.16385/j.cnki.issn.1004-4523.2024.03.013

引言

滚动轴承是旋转机械中的重要部件, 对其进行状态监测对于机械设备的正常运转至关重要。但采集到的滚动轴承振动信号往往包含大量的噪声, 对故障特征提取造成了困难^[1]。一般而言, 传感器采集的机械振动信号越多, 越有利于故障特征提取, 但这也给数据存储带来巨大的压力。因此, 在保留故障特征的前提下对振动响应信号进行高倍压缩意义重大。近年来, 从稀疏分解基础上发展而来的压缩感知(Compressed Sensing, CS)^[2-4]可实现信号的压缩, 并能利用低维观测信号重构高维原始信号, 为减轻数据存储和传输负担提供了建模思路。

目前, 压缩感知理论在机械故障诊断领域中已有一定的应用。Chen 等^[5]提出了一种基于字典学习和贪婪算法的压缩感知方法, 成功应用于齿轮箱的冲击特征提取。孟宗等^[6]基于压缩感知理论, 提出了一种自适应的分块匹配追踪算法, 有效提高了滚动轴承振动信号的重构效果。Wang 等^[7]对快速迭代收缩阈值算法进行了改进, 有效提高了压缩感知的重构精度, 并诊断出齿轮磨损故障。王强等^[8]

基于信号能量对信号进行分段, 利用学习字典和基追踪算法重构出轴承的故障特征。Lin 等^[9]利用移不变字典和多次压缩匹配追踪算法, 有效提取出了齿轮箱的冲击故障特征。Li 等^[10]对滚动轴承压缩数据进行独立成分分析, 利用重构的信号进行复合故障诊断。林慧斌等^[11]利用压缩信号的特征代理和凸优化算法对轴承故障信号进行有效重构。

以上文献在采用压缩感知进行信号重构时都涉及最优化解的迭代, 难以进行实时重构^[12]; 此外, 受噪声的影响, 传统方法利用测量矩阵对信号进行线性观测时, 在保证足够重构精度的前提下无法对信号进行大幅压缩。随着人工智能的快速发展, 一些学者将深度学习引入压缩感知中用以解决上述问题。Mousavi 等^[13]最先利用深度学习研究方法研究压缩感知, 提出了一种基于堆叠去噪自编码的信号压缩和重构方法。Jiang 等^[14]将卷积神经网络引入图片的压缩感知, 提出了一种端到端的压缩感知框架, 提高了图片的重构精度。Yao 等^[15]提出了一种深度残差重建网络, 进一步提高了图片的重构精度。Shi 等^[16]利用卷积层模拟传统压缩感知中的观测矩阵, 同时优化压缩端与重构端, 并将该方法成功应用于图片的压缩与重构, 并取得了比传统方法更快的重

收稿日期: 2022-05-31; 修订日期: 2022-07-25

基金项目: 国家自然科学基金资助项目(51875207); 广东省自然科学基金资助项目(2020A1515010750, 2022A1515011238)。

构速度。Metzler等^[17]基于传统压缩感知中的近似消息传递算法,设计了一种网络结构用于图像重构。Yang等^[18]利用神经网络将压缩感知中常用的交替方向乘子算法展开,并从随机测量值中更快地重构出原始图片。Yang等^[19]将生成模型与压缩感知结合,取得了优于传统方法的图像重构效率与精度。Cui等^[20]提出了一种基于非局部神经网络的深度压缩感知框架,提高了图片的重构质量。

从现有文献来看,基于深度学习的压缩感知在图像领域已取得一定的成效,但受机械振动复杂噪声的影响,鲜有在故障诊断领域中的应用,且现有压缩感知相关的文献中少有从信号压缩域中直接提取故障特征以完成故障诊断的方法。针对上述问题,本文结合深度学习与压缩感知理论,提出了一种基于深度卷积测量网络(Deep Convolutional Measurement Network, DCMN)的滚动轴承压缩域故障特征提取方法。提出了一种基于故障机理的数据集构造方法。该方法根据实际钢结构常见的固有频率和阻尼比取值范围,按照一定步长设置冲击响应模式生成数据集,能够解决实际工程中难以采集大量故障信号(特别是无噪声信号)用于模型训练的问题,而且采用所提数据集训练得到的网络具有很好的泛化性,适用于不同转速工况下冲击型故障的特征提取。此外,本研究利用自编码网络与压缩感知之间的对应关系,提出一种能够同时对信号进行压缩和降噪的深度卷积测量网络。所提网络利用卷积层和池化层分别实现特征提取和信号压缩,利用池化层的数量和池化域的宽度调整压缩率。基于该测量网络得到的压缩域信号无需重构便可以直接进行故障特征提取,且能达到比传统压缩感知方法更低的压缩率。该方法能够克服传统压缩感知理论在保证故障特征重构效果的前提下,无法大幅降低压缩率且压缩重构计算复杂度高的缺点。

1 理论基础

1.1 经典压缩感知理论

根据压缩感知理论^[3]可知,若信号本身具备稀疏性或在某一变换域 $\Psi \in \mathbf{R}^{N \times N}$ 下是稀疏的,则能够利用与变换域不相关的测量矩阵 $\Phi \in \mathbf{R}^{M \times N}$ ($M \ll N$)对该信号进行线性投影压缩,且可以通过优化算法从其压缩信号高概率地重构回原始信号。一般情况下传感器获得的信号 $x \in \mathbf{R}^{N \times 1}$ 都包含一定成分的噪声,即 $x = x_0 + e$,其中 x_0 表示无噪声的信号, e 代表噪声分量,则压缩信号 $y \in \mathbf{R}^{M \times 1}$ 可表示为:

$$y = \Phi(x_0 + e) = \Phi\Psi\theta + \Phi e \quad (1)$$

式中 θ 表示稀疏系数,为少量非零值组成的稀疏向量,非零值的个数称为稀疏度。由观测信号重构得到无噪声信号的过程可通过求解下式实现:

$$\min \|\theta\|_0, \quad \text{s.t.} \|\mathbf{y} - \Phi\Psi\theta\|_2^2 \leq \epsilon \quad (2)$$

式中 ϵ 为由噪声决定的误差限; $\|\cdot\|_0$ 为向量的 L_0 范数; $\|\cdot\|_2^2$ 为向量的 L_2 范数的二次方。

在压缩感知过程中,压缩率作为衡量压缩采集信号量的参数,一般定义为:

$$\Delta = \frac{M}{N} \times 100\% \quad (3)$$

式中 Δ 为压缩率; N 与 M 分别为原始信号和压缩信号的维度。 Δ 反映了对原始信号的压缩程度,即 Δ 越小压缩域数据越少,压缩程度越大。

1.2 自编码网络与压缩感知的关系

Hinton等^[21]提出的自编码网络是一种无监督神经网络,常用于数据的降维与特征学习,其基本架构如图1所示。图中,输入信号 x 经过编码函数 $g_e(\cdot)$ 后得到其压缩的隐含层信号 y ,而解码函数 $g_d(\cdot)$ 则是对隐含层信号 y 进行重构。通过最小化重构信号 x' 与输入信号 x 之间的均方误差(Mean Squared Error, MSE)完成模型的训练过程。

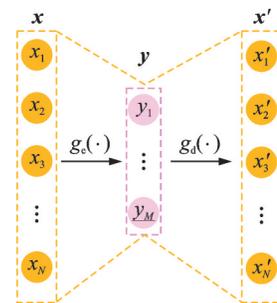


图1 自编码网络架构

Fig. 1 Architecture of the autoencoder network

通过分析压缩感知与自编码网络的原理可知,自编码网络的编码子网络将高维数据映射为低维数据,刚好对应压缩感知的信号观测过程,可将其称为测量网络;解码子网络将低维数据重构回原始高维数据,刚好与压缩感知中的信号重构过程相对应,可将其称为重构网络。压缩感知与自编码网络的对应关系如图2所示。

在自编码网络基础上发展起来的卷积自编码网络^[22]和去噪自编码网络^[23-24]同样与压缩感知具备上述对应关系。其中,卷积自编码网络将自编码网络引入卷积神经网络架构以提高模型的特征提取效果;去噪自编码网络以加噪后的信号作为输入、无噪

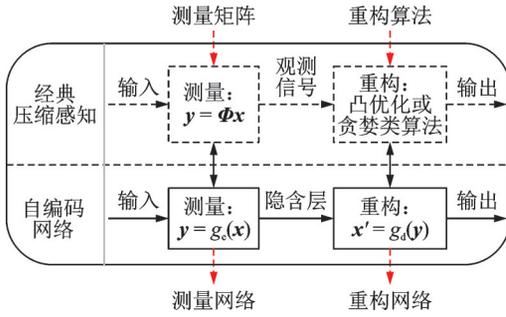


图2 压缩感知理论与自编码网络的对应关系

Fig. 2 The corresponding relationship between compressed sensing theory and autoencoder network

信号作为序列标注训练模型,使得模型具备一定的去噪效果。

2 基于深度卷积测量网络的压缩域故障特征提取

从上节内容可知,理论上可以采用训练好的编码网络代替观测矩阵对信号进行非线性测量,得到压缩观测信号,再利用训练好的解码网络代替传统的重构算法恢复原始信号。由于机械故障诊断的关键是故障特征的提取,若能根据需要合理地构造和训练用于代替线性测量矩阵的编码网络,使压缩信号保留原始故障特征,则可在不进行信号重构的情况下进行故障诊断。基于上述思想,提出了一种基于深度卷积测量网络(DCMN)的压缩域故障特征提取方法。

2.1 基于故障机理的数据集构造方法

为了使测量网络能在对信号进行压缩的同时有效地保留故障特征并减少噪声的影响,去噪自编码网络需要大量的含噪信号和与之对应的无噪信号作为训练样本,但实际工程中采集到的机械振动信号都包含一定的噪声,无法满足去噪自编码网络模型训练的样本需求。为解决上述问题,提出基于故障机理的数据集构造方法。

当滚动轴承发生局部故障时,故障表面会因与其他表面碰撞而产生冲击力,导致系统发生结构共振^[25]。故障激起的冲击响应信号可由下面两式给出^[26]:

$$x(t) = x_o(t) + e(t) = \sum_i A_i h(t - iT) + e(t) \quad (4)$$

$$h(t) = \exp\left(\frac{-2\pi\xi}{\sqrt{1-\xi^2}} f_d t\right) \cdot \sin(2\pi f_d t) \quad (5)$$

式中 $x(t)$ 表示含噪的故障响应信号; $x_o(t)$ 表示故障激起的冲击成分; A_i 表示第 i 个冲击响应的振幅;

T 表示冲击时间间隔; $e(t)$ 为均值为 0 的高斯白噪声; $h(t)$ 表示故障冲击响应; ξ 和 f_d 分别为阻尼比和固有频率。

由上述滚动轴承局部故障信号数学模型可知,故障冲击分量 $x_o(t)$ 还可以表示为故障冲击响应 $h(t)$ 与包含冲击响应发生时刻和冲击幅值信息的稀疏系数 θ 的卷积,即:

$$x_o(t) = h(t) * \theta \quad (6)$$

因此,可以根据式(4)和式(6)分别生成含有噪声信号和对应无噪声信号的样本,用于去噪自编码网络的参数训练。

为了提高数据集的适用范围,样本的固有频率和阻尼比可根据实际工程中钢结构的常见固有频率和阻尼比的范围^[27]按一定步长进行取值,按式(5)生成一组故障冲击响应模式。为了生成大量样本,可利用计算机软件(如 Scikit-learn)生成足够多幅值随机取值、位置随机出现、稀疏度为 1 的稀疏系数 θ 。将上述生成的冲击模式与稀疏系数按式(6)进行卷积运算,即可得到大量只包含一个冲击模式的无噪信号 $x_o(t)$ 样本。再对生成的无噪信号 $x_o(t)$ 添加高斯白噪声,便可得到相应的含噪冲击信号 $x(t)$ 样本。将生成的含噪样本与无噪样本的幅值进行归一化处理,以含噪信号样本作为输入,无噪信号样本作为序列标注,完成数据集的构建。

这种数据集构造方法具有如下优点:(1)该方法基于故障机理生成数据集,利用该数据集训练得到的网络适用于具有相同故障机理的局部故障信号的特征提取,解决了实际工程中难以获得大量训练样本的问题。(2)由于含噪样本及其相应的无噪信号均已知,以含噪信号为输入,无噪信号为序列标注,训练得到的网络具有很好的降噪效果。(3)数据集的每个样本只包含一个冲击模式,而冲击位置是随机生成的,故采用该数据集训练得到的网络适用于不同转速工况下的故障冲击提取。

2.2 深度卷积测量网络的设计方法

由 1.2 节的分析可知,传统压缩感知的信号观测和压缩信号重建与自编码网络的编码过程和解码过程存在对应关系。本节通过引入深度卷积神经网络^[28],构造包含卷积层、池化层和上采样层的深度卷积去噪自编码(Deep Convolutional Denoising Auto-encoder, DCDAE)网络,再利用基于 2.1 节的方法生成的数据集进行训练,然后截取该网络的编码部分,得到能够在压缩域进行故障特征提取的深度卷积测量网络(DCMN)。在 DCDAE 网络构建过程中,为

了使编码得到的隐含层信号满足预先设定的压缩率要求,同时便于建立编码后的压缩域信号与原信号之间的频率对应关系,所提网络选用最大池化层作为信号压缩的手段,而卷积层仅用作冲击特征提取,不改变信号的维度。所提深度卷积去噪自编码网络具体设计依据如下:

(1) 每个卷积层输出的维度 d' 等于输入的维度 d , 并依据下式合理设置卷积核的参数:

$$d' = \frac{d - h + 2p}{s} + 1 \quad (7)$$

式中 h 表示卷积核的大小; p 表示 Padding 的大小; s 表示 Stride 的大小。按此原则设置 DCDAE 网络的卷积层参数, 便于后续利用最大池化层建立起压缩域与原信号之间的频率对应关系。此外, 为提高网络的非线性学习能力, 常在卷积层后施加 ReLU 激活函数用于非线性变换。但 ReLU 激活函数会将负值置 0, 为使压缩域和重构信号保留负值特征, 在编码子网络和解码子网络的最后一层不施加激活函数。

(2) 最大池化层起到压缩信号的作用, 设置在编码子网络的卷积层后, 其原理为:

$$Q_i(g) = \max_{(j-1)W+1 \leq u \leq jW} \{Z_i(u)\} \quad (8)$$

式中 $Z_i(u)$ 为池化层第 i 个输入特征矢量中的第 u 个分量; W 为池化域的宽度; $Q_i(g)$ 为池化层第 i 个

输出特征矢量中的第 g 个分量; j 表示当前池化域位置。由式(8)可知, 最大池化层以 W 为间隔对输入数据进行抽样, 以达到压缩信号的效果。

(3) 卷积核数量会影响卷积层的输出数据量和特征提取效果, 为使得 DCDAE 网络的编码子网络的输出数据量少于输入, 编码子网络的卷积核数量以 2 的倍数逐层递减, 最后一层卷积核数量为 1, 从而达到利用池化层进行信号压缩的目的。其中, 第一层卷积用于初步故障特征提取, 应提取出足够多的特征供深层网络学习, 故第一层的卷积核数量 K 最多。后续的卷积层用于故障特征的筛选与深层特征的提取, 卷积核数量逐层递减。

(4) 网络的具体层数由实际所需的压缩率 Δ 决定, 编码子网络中每增加一层池化层, 网络就会对数据多压缩 W 倍, 相应地, 由于特征维度变少, 网络需要具备更强的特征提取能力, 因此每增加一层池化层, 也相应增加一层卷积层用于特征提取。

图 3 所示为所提的 DCDAE 网络的模型架构, 其中 N 为输入层的长度, M 为隐层信号的长度。编码子网络利用卷积层和池化层搭建, 解码子网络则利用卷积层和上采样层搭建, 与编码子网络采用对称结构。按此原则设置的网络会随着压缩率的降低而变深, 但由于卷积核数量逐层递减, 网络的参数量并不会随着压缩率的降低而大幅增加。

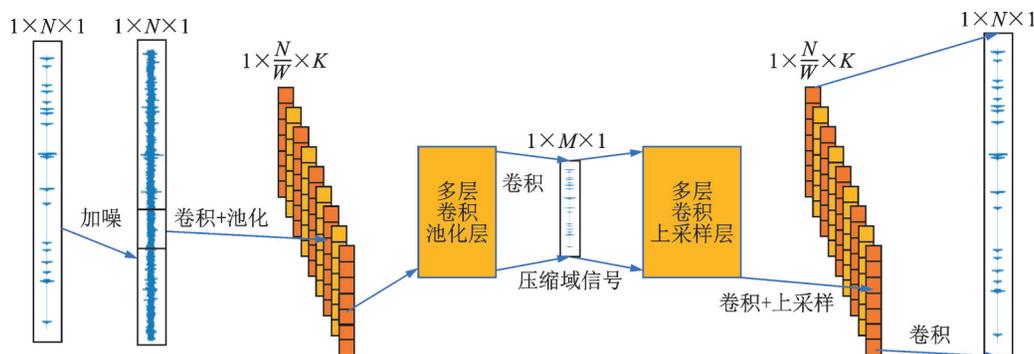


图 3 DCDAE 网络架构

Fig. 3 Architecture of the DCDAE network

搭建完 DCDAE 网络后, 利用 2.1 节所述的方法构造数据集对模型进行训练, 截取训练后的网络编码部分, 即得到本文所提的深度卷积测量网络 (DCMN)。

采用上述测量网络得到的隐含层的压缩率 Δ 完全由测量网络包含的池化层数量 r 及池化域的宽度 W 决定, 即 $\Delta = 1/W^r$ 。得到的压缩域信号 (即隐含层) 频率 f' 与原始信号频率 f 的关系为:

$$f = \Delta \times f' \quad (9)$$

图 4 所示为 $N=8192$, $M=512$, $\Delta=6.25\%$,

$W=2$, $K=128$ 时对应 DCMN 的具体架构。

表 1 给出了按上述原则确定的压缩率为 6.25% 时 DCDAE 网络的参数。

2.3 所提方法具体步骤

所提方法的基本思想是利用基于故障机理生成的数据集训练深度卷积去噪自编码网络, 截取其编码部分作为深度卷积测量网络代替传统压缩感知中的线性测量矩阵, 同时实现信号的压缩与去噪, 进而得到包含故障信息的压缩域信号, 最后从压缩域信

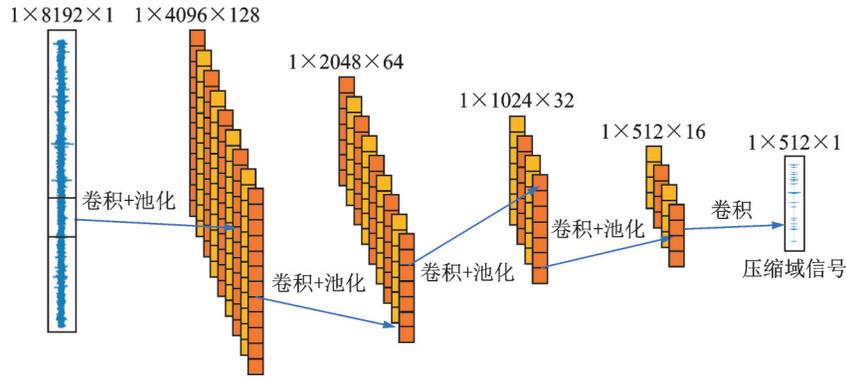


图 4 压缩率为 6.25% 时 DCMN 的架构

Fig. 4 Architecture of DCMN with compression ratio of 6.25%

表 1 压缩率为 6.25% 时 DCDAE 网络的参数

Tab. 1 DCDAE network parameters with compression ratio of 6.25%

层名	卷积核数量 K	卷积核大小 h	池化域宽度/ 上采样层倍率 W	激活函数
Conv1	128	5	2	ReLU
Conv2	64	5	2	ReLU
Conv3	32	5	2	ReLU
Conv4	16	5	2	ReLU
Conv5	1	5	—	—
Conv6	16	5	2	ReLU
Conv7	32	5	2	ReLU
Conv8	64	5	2	ReLU
Conv9	128	5	2	ReLU
Conv10	1	5	—	—

号中完成故障特征提取与诊断。所提方法的一般步骤如下:

(1) 基于滚动轴承故障机理,以 2.1 节所提方法生成数据集。

(2) 利用自编码网络与压缩感知之间的对应关系,按照预期的压缩率,按 2.2 节所提方法构建 DCDAE 网络模型,并采用步骤(1)中生成的数据集对模型进行训练。

(3) 截取训练完备的 DCDAE 网络的编码子网络得到本文所提 DCMN。

(4) 对采集到的滚动轴承振动信号进行归一化处理,并利用 DCMN 进行非线性测量,得到压缩后的信号。

(5) 对压缩域信号直接进行希尔伯特解调分析,并利用式(9)进行频率换算,获取故障特征频率,完成故障诊断。

具体流程图如图 5 所示。

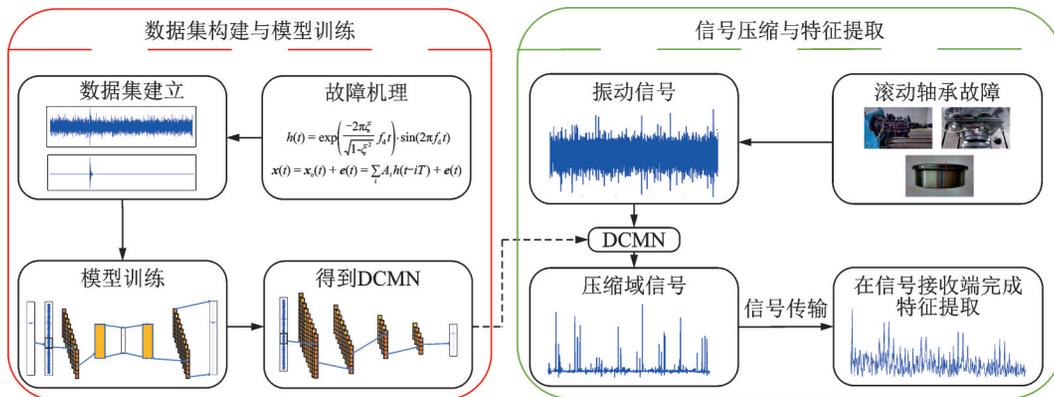


图 5 所提方法的流程图

Fig. 5 Flow chart of the proposed method

3 参数选取策略与仿真实证

按 2.1 节所述的方法构造数据集,一般而言,数据集的参数取值间隔越密,训练样本越多,训练效果就越好,但训练耗时也越长。综合考虑网络训练的速度和效果,本文所用的训练集的固有频率 f_0 在

1k~10k Hz 范围内,以 500 Hz 为步长。考虑到阻尼比主要影响冲击衰减速度,对故障特征影响较小(实测发现数据集阻尼比在 0.01~0.2 范围内取值对故障特征提取结果影响不明显),故将本文的仿真数据集阻尼比 ξ 设置为定值 0.05。为使得训练集更贴近于真实冲击,数据集中的冲击幅值在 0.1~1 m/s² 之间随机取值,以使样本包含不同冲击幅值的情形。

由于所构造网络以 2 的倍数对信号进行压缩, 此处样本的分析点数设置为 2 的指数倍 8192。考虑到实际测试时采样频率一般按分析频率的几倍频进行设置, 每个样本的采样频率 f_s 与固有频率 f_d 的比值 λ 在 [4, 10] 内随机赋一个值。由于数据集中每个样本仅设置单个冲击, 若要加入足够的噪声分量, 需要将信噪比 (Signal to Noise Ratio, SNR) 设置为较低值。按上述参数生成无噪冲击样本, 并考虑到噪声向上兼容原则, 含噪样本的信噪比设为 -20 dB, 归一化后得到包含 10000 个无噪和有噪样本的数据集。以 0.8:0.2 的比例将数据集划分为训练集和验证集, 其中训练集用于对按照 2.2 节所述方法构造的深度卷积去噪自编码网络进行训练, 验证集用于在训练过程中验证模型的泛化性能, 防止模型过拟合, 训练的 epoch 设置为 50。对训练好的 DCDAE 网络截取其编码部分, 得到网络结构如表 1 所述的训练好的 DCMN。本文下面的仿真和实验分析均基于上述训练好的网络。

3.1 数据集构造方法有效性验证

为了验证上述数据集的有效性, 本节按式 (4) 生成信噪比 SNR 分别为 0, -5 和 -8 dB 的滚动轴承局部故障仿真信号, 采样频率设置为 25600 Hz, 采样时长为 0.32 s (即 8192 点), 冲击周期 $T=0.02$ s (即共包含 16 个理论冲击), 冲击幅值在 $0.5\sim 1.5$ m/s^2 之间随机分布。由于故障冲击的间隔是最重要的故障特征参数, 将利用上述训练好的 DCDAE 网络重构得到的降噪后的信号与原始无噪信号进行对比, 将在正确位置重构出冲击响应的数量与理论冲击数量之比 α 作为重构精度的评价指标, 讨论本文所提数据集构造方法的有效性。显然 $\alpha \in [0, 1]$, 其值越接近 1, 重构的效率越高。

为验证构造的数据集是否适用于不同的固有频率, 该组仿真信号的阻尼比设置为定值 0.05, 固有频率在 2560~6400 Hz 范围内取 13 个值, 使比值 λ 以 0.5 为步长落在 [4, 10] 内。将采用上述参数生成的仿真信号通过所提的压缩率为 6.25% 的 DCDAE 网络进行故障冲击提取, 不同信噪比及 λ 值下的重构精度如图 6 所示 (为了减少随机噪声带来的不确定性的影响, 图中各点均是 200 次重复测试的平均值)。由图可知, 冲击信号的重构精度在不同噪声水平下均变化不大, 而且即使在 -8 dB 的低信噪比下, 仍有 75% 以上的冲击位置可以被准确地重构, 验证了所提数据集参数设置方法对不同固有频率的有效性。

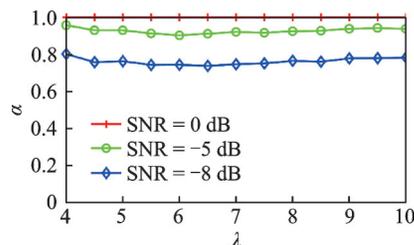


图 6 不同信噪比及 λ 值下的重构精度

Fig. 6 Reconstruction accuracy under different signal to noise ratio and λ

为讨论阻尼比对所提数据集有效性的影响, 仿真信号的采样频率、采样时长、冲击周期和冲击幅值设置同上, 取固有频率为 2560 Hz, 阻尼比则以 0.01 为步长在 0.04~0.1 内选取。将采用上述参数生成的仿真信号通过上述相同的 DCDAE 网络进行故障冲击提取, 每种工况重复 200 次, 得到的不同信噪比及阻尼比下的平均重构精度如图 7 所示。由图 7 可见, 不同噪声水平下冲击信号的重构精度并没有因为仿真信号的阻尼比与数据集设置的阻尼比不同而产生明显变化, 且即使是在 -8 dB 的低信噪比下仍有接近 80% 的冲击位置可以被准确地重构。

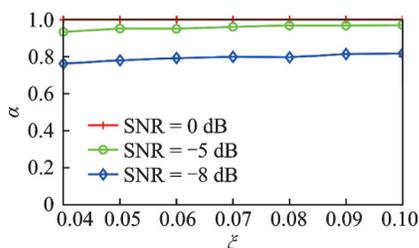


图 7 不同信噪比及阻尼比下的重构精度

Fig. 7 Reconstruction accuracy under different signal to noise ratio and ξ

通过以上仿真分析可知, 本文所提的数据集的构造方法对具有不同固有频率和阻尼比的冲击特征的提取都是有效的, 特别是在信噪比为 0 和 -5 dB 的情况下可以达到很高的重构精度。

3.2 压缩域故障特征提取结果分析

为了验证所提的深度卷积测量网络可以在压缩域直接进行故障特征提取, 同样利用式 (4) 生成信噪比为 -9 dB, 包含 16 个冲击的仿真信号。其中采样频率 $f_s=25600$ Hz, 固有频率 $f_d=3000$ Hz, 阻尼比 $\xi=0.06$, 冲击周期 $T=0.02$ s, 冲击幅值在 $0.5\sim 1.5$ m/s^2 之间随机分布, 采样时长为 0.32 s。所得仿真信号的时域波形如图 8 所示, 由图可见, 此时故障冲击被强噪声淹没, 无法从时域中辨别。

将该信号幅值进行归一化处理, 再输入到按上述方法训练好的深度卷积测量网络中, 得到压缩率

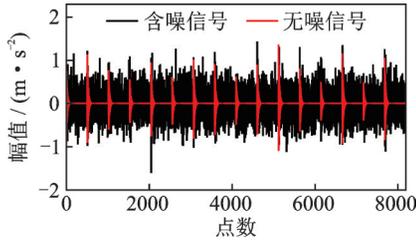
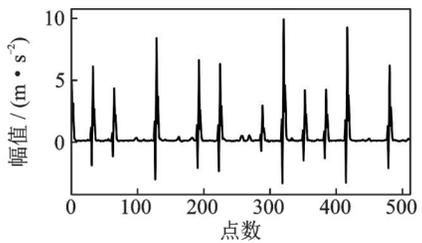


图 8 仿真信号时域波形

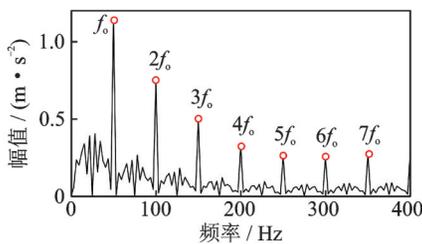
Fig. 8 Time domain waveform of simulation signal

为 6.25% 的压缩域信号如图 9(a) 所示。将压缩域信号直接进行希尔伯特解调处理。得到频率按式 (9) 还原后的解调谱如图 9(b) 所示。

从图 9(a) 可知, 尽管丢失了原冲击衰减信号的固有频率和阻尼比信息, 按本文所提方法构造的深度卷积测量网络得到的压缩域信号保留了原信号故障冲击间隔的特征, 且去除了绝大部分的噪声分量。图 9(b) 所示的压缩域信号解调谱中前 7 阶故障特征频率均清晰可见, 且各阶幅值呈现有规律的递减, 表明所提压缩域故障特征提取方法是可行的。



(a) 压缩域信号
(a) Compressed domain signal



(b) 压缩域信号解调谱
(b) Demodulation spectrum of the compressed domain signal

图 9 压缩域信号及解调谱

Fig. 9 Compressed domain signal and demodulation spectrum

3.3 不同压缩率下模型性能对比分析

仍用 3.2 节定义的信噪比为 -9 dB 的仿真信号进一步分析所提模型在不同压缩率下的性能。受所提网络结构参数的影响, 所提模型只能以 2 为倍数对信号进行压缩。表 2 对不同压缩率下的模型进行了对比, 训练时间为得到不同压缩率的测量网络的时间, 重构精度仍用 3.1 节定义的在正确位置重构出冲击响应的数量与理论冲击数量之比 α 来衡量, 运算时间指采用训练好的测量网络处理图 8 所示信

号时的运算时间。用于运算的计算机配置如下: CPU 为 AMD R7 5800H, GPU 为 RTX3060, 运行内存为 16 GB。

表 2 不同压缩率下模型对比

Tab. 2 Comparison of models with different compression rates

压缩率/%	训练时间/s	重构精度 α	运算时间/s
12.5	815	0.75	0.0011
6.25	890	0.75	0.0013
3.125	908	0.75	0.0016
1.5625	920	0.75	0.0017

对表 2 分析可知, 随着压缩率从 12.5% 降低到 1.5625%, 所提 DCMN 网络层数增加, 模型的重构精度 α 并没有降低, 在压缩率低至 1.5625% (即压缩 64 倍) 时故障冲击在正确位置上的重构率仍可达到 75%。此外还可以发现, 随着压缩率的降低, 网络层数虽然在增加, 但由于网络深层卷积核数量和特征量均较少, 按所提方法构造的模型的训练时间和压缩特征运算时间并未显著增加。

进一步分析发现, 当压缩率进一步降低时, 所提网络会存在不收敛的现象, 这是因为当原始信号被压缩 128 倍时, 在压缩域仅有极少数点来表征一个冲击, 由于特征量过少导致解码网络难以完成重构, 以致网络难以收敛。

为进一步验证所提方法的有效性, 采用常见的压缩感知方法与所提方法进行对比。对比方法的测量矩阵选取常用的高斯随机矩阵, 稀疏字典及其参数选取方法参考文献 [7] 提出的移不变 K-SVD (Shift Invariant K-Singular Value Decomposition) 字典, 重构算法选用压缩感知中常用的 CoSaMP 算法^[29], 由于 CoSaMP 算法进行压缩重构时需预估原信号的稀疏度, 此处直接将稀疏度设置为理论值 16。表 3 为压缩率取 12.5% 与 6.25% 时对比方法的重构精度及运算时间。

由表 3 可知, 在 -9 dB 的噪声水平下, 对比方法在压缩率为 12.5% 时只能重构出半数的冲击, 极易发生误诊; 而在压缩率为 6.25% 时仅有 30% 左右的冲击能在正确位置被重构, 无法用于诊断, 该结果也和文献 [9] 中在采用传统压缩感知进行机械故障诊断时压缩率普遍不低于 20% 的结论一致。

表 3 对比方法性能分析

Tab. 3 Performance analysis of comparison method

压缩率/%	重构精度 α	运算时间/s
12.5	0.5	0.079
6.25	0.3125	0.064

此外,对比表2和3可知,所提方法在压缩率为12.5%和6.25%时的运算时间仅为对比方法的1.3%和2%,显著低于对比方法。表明本文所提的方法不仅可以达到远低于传统压缩感知的压缩率,且能大幅提高故障诊断的精度和效率,更适合用于机械故障信号的压缩传输和在线诊断。

4 实验信号分析

为进一步验证所提数据集构造方法的适用性以及压缩域故障特征重构方法的有效性,采用上述经仿真信号训练的深度卷积测量网络对实测的轴承故障信号进行压缩和特征提取。

4.1 滚动轴承外圈实验信号分析

将N205M圆柱滚子轴承作为被测对象,在图10所示的实验台上进行实验,安装在轴承座上的振动加速度传感器以 $f_s=100$ kHz的采样频率采集振动信号。外圈故障以线切割方式进行加工,故障深度为1 mm,宽度为0.5 mm。将故障轴承所在轴的转速设置为500 r/min,则计算得到其转频为 $f_n=8.33$ Hz。根据轴承的结构参数可得到轴承故障特征频率^[30],计算得到轴承外、内圈的故障特征频率分别为 $f_o=44.9$ Hz和 $f_i=63.43$ Hz。



图10 实验台与故障外圈

Fig. 10 The test bench and rolling bearing with outer race fault

从采集到的振动信号中截取时长为1 s的信号进行分析,所得信号时域波形、频谱及其希尔伯特解调谱如图11所示,由图可见,受强噪声干扰,时域信号并无明显的冲击成分,频谱中并无明显的共振峰,而解调谱中虽然也出现了故障特征频率 f_o 的2倍及4倍频,但这两个故障频率成分在解调谱中均不明显,该解调谱无法用于诊断。

将信号以8192点进行分段,归一化后依次输入到第3节中经仿真信号训练的压缩率为6.25%的DCMN中,再将压缩后的信号进行简单拼接,得到的压缩域信号如图12(a)所示。由图可知,得到的压缩域信号中出现了较明显的有规律冲击成分,而原始信号中的大部分噪声被去除。将得到的压缩域

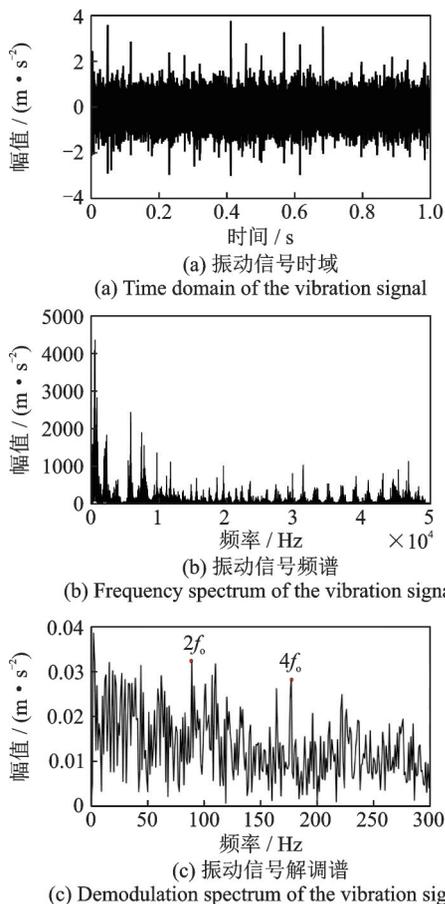


图11 外圈故障振动信号的时域、频谱及解调谱

Fig. 11 Time domain, frequency spectrum and demodulation spectrum of the vibration signal with outer race fault

信号直接进行希尔伯特解调处理,并按式(9)进行频率还原,得到压缩域信号的解调谱如图12(b)所示。由图可知,此时压缩域信号的解调谱中出现了明显的外圈故障特征频率及其倍频,由此可以直接诊断该轴承存在外圈故障。

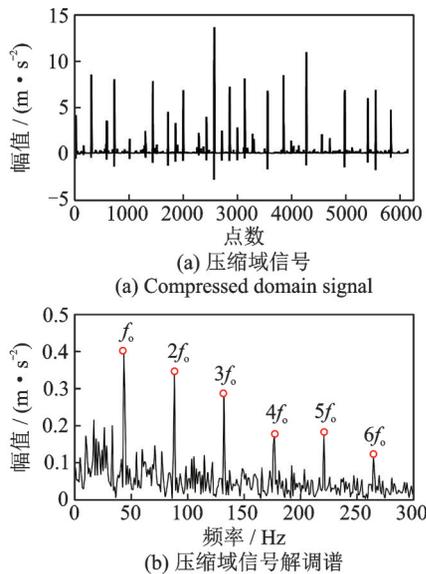


图12 压缩率为6.25%时的压缩域信号及解调谱

Fig. 12 Compressed domain signal and demodulation spectrum with compression ratio of 6.25%

同样的,将信号输入到训练好的压缩率为 1.5625% 的 DCMN 中,得到其希尔伯特解调谱如图 13 所示。显然,即使是在压缩率为 1.5625% 的情况下,所提方法仍可以有效地提取故障特征。

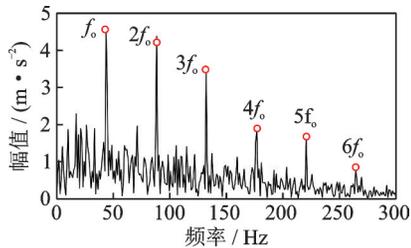


图 13 压缩率为 1.5625% 时的压缩域信号解调谱

Fig. 13 Demodulation spectrum of compressed domain signal with compression ratio of 1.5625%

以 3.3 节中提到的传统压缩感知方法进行对比,采用高斯随机矩阵以 6.25% 的压缩率进行观测,采用 CoSaMP 算法得到的重构信号及其希尔伯特解调谱如图 14 所示。由图可见,即使是在给出理论稀疏度的情况下,传统压缩感知方法在 6.25% 的压缩率下的重构性能仍不佳,重构信号的解调谱中包括大量的噪声,基本已经无法用于诊断。

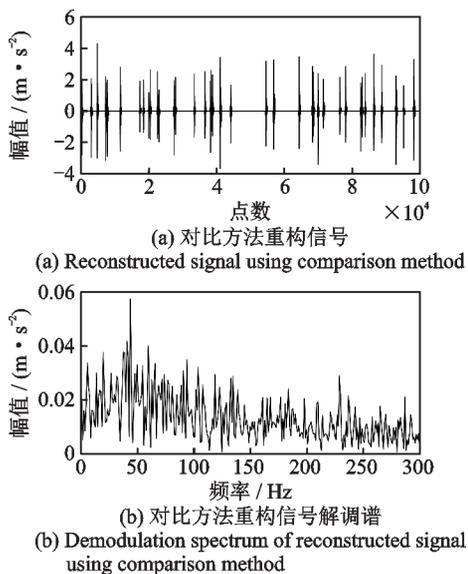


图 14 对比方法外圈故障信号重构结果

Fig. 14 Reconstruction results using comparison method of the vibration signal with outer race fault

4.2 滚动轴承内圈实验信号分析

进一步以安置在三轴五档变速器输出轴的 NUP311EN 圆柱滚子轴承为研究对象,内圈故障采用线切割加工,故障尺寸为 1 mm 深,0.2 mm 宽,加速度传感器安置在接近轴承的壳体上,实验所用变速器、传感器安装位置及故障内圈如图 15 所示。实验过程中,输入轴转速设置为 1500 r/min,施加

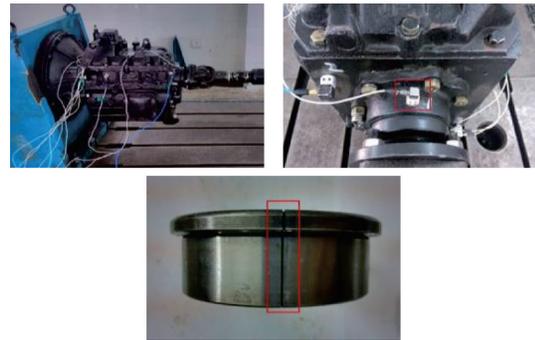


图 15 实验变速器、传感器安装位置及内圈故障

Fig. 15 The tested transmission, sensor's location and inner race fault

45 N·m 的负载,变速器置于二档,根据传动比可计算得到输出轴转频 $f_n=8.34$ Hz,同理可得到轴承的外、内圈的特征频率分别为 $f_o=42.73$ Hz 和 $f_i=65.69$ Hz。振动信号的采样频率取 $f_s=100$ kHz。

同样截取时长为 1 s 的信号进行分析,原始振动信号的时域波形、频谱及其解调谱如图 16 所示,受噪声影响,图 16 中由故障激起的冲击成分并不明显,虽然图 16(c) 的解调谱中有内圈故障特征频率 f_i 的 2 倍频,但是转频调制边带不明显,无法直接判断该变速器出现哪类故障。

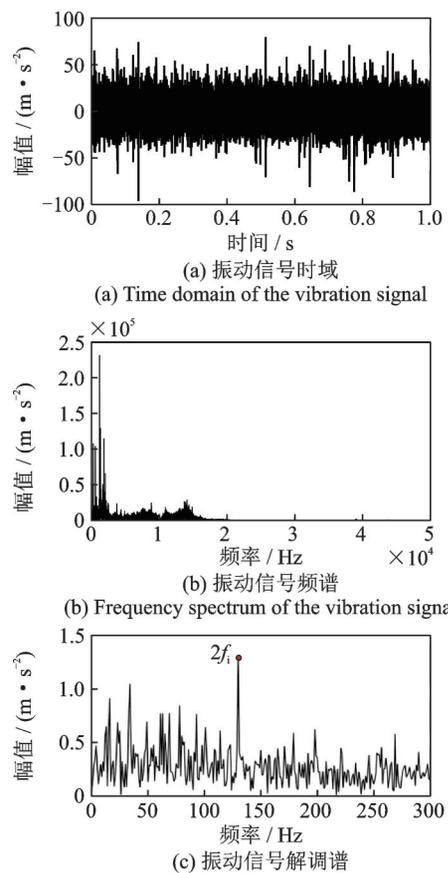
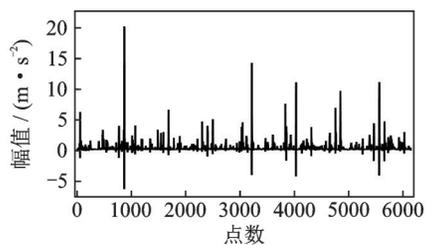


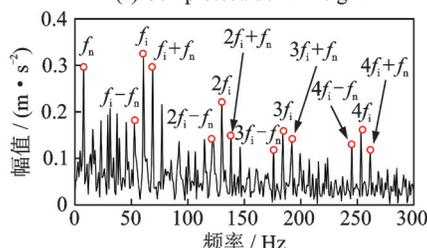
图 16 内圈故障振动信号的时域、频谱及解调谱

Fig. 16 Time domain, frequency spectrum and demodulation spectrum of the vibration signal with inner race fault

将振动信号以 8192 点进行分段,归一化后依次输入到上述由仿真信号训练好的 DCMN 中,再对拼接得到的压缩信号进行解调和频率还原,在压缩率为 6.25% 时得到的结果如图 17 所示。由图可见,所提方法压缩域信号很好地保留了冲击的位置信息,其解调谱中的转频 f_n ,内圈故障特征频率 f_i 的前 4 倍频以及相应的调制边带均可清晰分辨,且幅值呈现有规律的递减趋势,可以直观诊断出轴承内圈发生故障。



(a) 压缩域信号
(a) Compressed domain signal



(b) 压缩域信号解调谱
(b) Demodulation spectrum of compressed domain signal

图 17 压缩率为 6.25% 时压缩域信号及解调谱

Fig. 17 Compressed domain signal and demodulation spectrum with compression ratio of 6.25%

将故障信号用上述压缩率为 1.5625% 的测量网络进行压缩,得到的压缩域信号及其解调谱如图 18 所示。同样可以看出,采用本文所提方法,即使是在 1.5625% 的低压缩率下仍旧有非常好的故障特征提取效果。

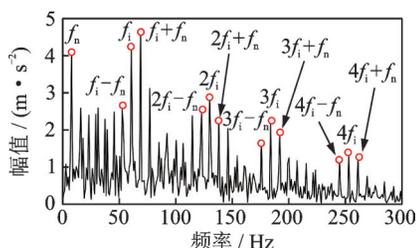
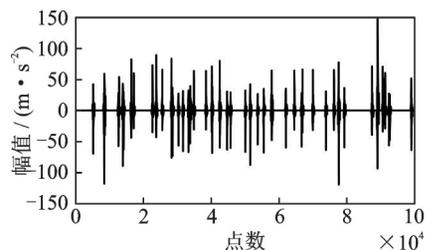


图 18 压缩率为 1.5625% 时的压缩域信号解调谱

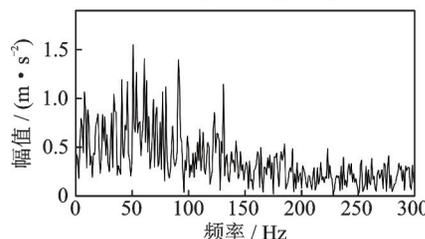
Fig. 18 Demodulation spectrum of compressed domain signal with compression ratio of 1.5625%

以 3.3 中提到的传统压缩感知方法作为对比,并将对比方法的稀疏度设为理论值,对比方法在压缩率为 6.25% 时的重构信号及其解调谱如图 19 所示。



(a) 对比方法重构信号

(a) Reconstructed signal using comparison method



(b) 对比方法重构信号解调谱

(b) Demodulation spectrum of reconstructed signal using comparison method

图 19 对比方法内圈故障信号重构结果

Fig. 19 Reconstruction results using comparison method of the vibration signal with inner race fault

如图 19 所示,在压缩率为 6.25% 时,传统的压缩感知方法重构的信号丢失了大量的故障特征信息,解调谱主要成分为未知的噪声分量,表明传统压缩感知方法在压缩率为 6.25% 时已无法有效用于故障诊断。相比之下,所提方法在低至 1.5625% 的压缩率下依旧能对故障特征进行有效提取,而且无需对压缩信号进行重构,本例也进一步验证了所提方法的有效性。

5 结 论

本文提出了一种基于深度卷积测量网络的压缩域故障特征提取方法。该方法利用滚动轴承局部故障振动信号模型构造数据集;利用自编码网络与压缩感知的对应关系建立深度卷积测量网络;利用训练完备的深度卷积测量网络代替传统的测量矩阵对振动信号进行观测,得到包含故障特征的压缩域信号,直接从压缩域提取出故障特征完成故障诊断。主要结论包括:

(1) 基于故障机理的数据集构造方法解决了实际工程中难以获取大量样本,特别是无噪声样本进行训练的问题。该数据集每个样本只包含一个冲击,有效性主要取决于仿真样本的采样频率与固有频率的比值,阻尼比影响较小。

(2) 所构造的深度卷积去噪自编码网络以卷积层进行特征提取,以池化层作为信号压缩的手段,利用预期压缩率调整池化层参数。该网络能设置的压缩率由测量网络包含的池化层数量及池化域的宽度

决定。所提网络在压缩率为1.5625%时仍可以取得很好的特征提取效果,但若进一步降低压缩率会出现网络不收敛的现象。

(3)采用所提数据集训练得到的深度卷积测量网络具有很好的泛化性,一次训练完毕后可以直接用于不同类型的实测轴承故障信号的压缩和故障特征提取。

(4)采用所提的深度卷积测量网络代替传统压缩感知方法的随机矩阵进行信号观测,可在大幅降低压缩率的同时有效保留故障特征,且在压缩域即可实现故障诊断。

参考文献:

- [1] 丁康,黄志东,林慧斌.一种谱峭度和Morlet小波的滚动轴承微弱故障诊断方法[J].振动工程学报,2014,27(1):128-135.
Ding Kang, Huang Zhidong, Lin Huibin. A weak fault diagnosis method for rolling bearings based on Morlet wavelet and spectral kurtosis [J]. Journal of Vibration Engineering, 2014, 27 (1): 128-135.
- [2] Candès E J, Romberg J K, Tao T. Stable signal recovery from incomplete and inaccurate measurements [J]. Communications on Pure and Applied Mathematics, 2006, 59(8): 1207-1223.
- [3] Donoho D L. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [4] Candès E J, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information [J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.
- [5] Chen X F, Du Z H, Li J M, et al. Compressed sensing based on dictionary learning for extracting impulse components [J]. Signal Processing, 2014, 96(5): 94-109.
- [6] 孟宗,石颖,潘作舟,等.自适应分块前向后向分段正交匹配追踪在重构滚动轴承故障信号中应用[J].机械工程学报,2020,56(9):91-101.
Meng Zong, Shi Ying, Pan Zuozhou, et al. Fault diagnosis of rolling bearing based on adaptive block forward and backward stagewise orthogonal matching pursuit algorithm [J]. Journal of Mechanical Engineering, 2020, 56(9): 91-101.
- [7] Wang Q, Meng C, Ma W N, et al. Compressive sensing reconstruction for vibration signals based on the improved fast iterative shrinkage-thresholding algorithm [J]. Measurement, 2019, 142: 68-78.
- [8] 王强,张培林,王怀光,等.基于优化分类的机械振动信号压缩感知[J].振动与冲击,2018,37(14):86-93.
Wang Qiang, Zhang Peilin, Wang Huaiguang, et al. Compressed sensing algorithm for machinery vibration signals based on optimal classification [J]. Journal of Vibration and Shock, 2018, 37(14): 86-93.
- [9] Lin H B, Tang J M, Mechefske C. Impulse detection using a shift-invariant dictionary and multiple compressions [J]. Journal of Sound and Vibration, 2019, 449: 1-17.
- [10] Li Jing, Meng Zong, Yin Na, et al. Multi-source feature extraction of rolling bearing compression measurement signal based on independent component analysis [J]. Measurement, 2021, 172: 108908.
- [11] 林慧斌,邓立发.滚动轴承压缩故障信号的特征代理与凸优化重构算法 [J]. 振动工程学报, 2022, 35(2): 434-445.
Lin Huibin, Deng Lifa. Feature proxy and convex optimization reconstruction algorithm for rolling bearing compressed fault signal [J]. Journal of Vibration Engineering, 2022, 35(2): 434-445.
- [12] Metzler C A, Maleki A, Baraniuk R G. From denoising to compressed sensing [J]. IEEE Transactions on Information Theory, 2016, 62(9): 5117-5144.
- [13] Mousavi A, Patel A B, Baraniuk R G. A deep learning approach to structured signal recovery [C]//2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton). IEEE, 2015: 1336-1343.
- [14] Jiang F, Tao W, Liu S H, et al. An end-to-end compression framework based on convolutional neural networks [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 28(10): 3007-3018.
- [15] Yao H T, Dai F, Zhang S L, et al. DR²-Net: deep residual reconstruction network for image compressive sensing [J]. Neurocomputing, 2019, 359: 483-493.
- [16] Shi W Z, Jiang F, Zhang S P, et al. Deep networks for compressed image sensing [C]//Proceedings of the IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2017: 877-882.
- [17] Metzler C A, Mousavi A, Baraniuk R G. Learned D-AMP: principled neural network based compressive image recovery [C]//Proceedings of Advances in Neural Information Processing Systems (NIPS 2017). 2017: 1773-1784.
- [18] Yang Y, Sun J, Li H B, et al. ADMM-CSNet: a deep learning approach for image compressive sensing [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(3): 521-538.
- [19] Yang G, Yu S M, Dong H, et al. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction [J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1310-1321.
- [20] Cui W X, Liu S H, Jiang F, et al. Image compressed sensing using non-local neural network [J]. IEEE Trans-

- actions on Multimedia, 2021, 25: 816-830.
- [21] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [22] Masci J, Meier U, Cireşan D, et al. Stacked convolutional auto-encoders for hierarchical feature extraction [C]//Proceedings of Artificial Neural Networks and Machine Learning-ICANN 2011. Springer, 2011: 52-59.
- [23] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]//Proceedings of the 25th International Conference on Machine Learning. New York, United States: Association for Computing Machinery, 2008: 1096-1103.
- [24] 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.
- [25] Antoni J, Randall R B. A stochastic model for simulation and diagnostics of rolling element bearings with localized faults[J]. Journal of Vibration and Acoustics, 2003, 125(3): 282-289.
- [26] Stevenson J D. Structural damping values as a function of dynamic response stress and deformation levels [J]. Nuclear Engineering and Design, 1980, 60(2): 211-237.
- [27] 丁玉兰, 石来德. 机械设备故障诊断技术[M]. 上海: 上海科学技术文献出版社, 1994.
- [28] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation, 1989, 1(4): 541-551.
- [29] Needell D, Tropp J A. CoSaMP: iterative signal recovery from incomplete and inaccurate samples[J]. Applied and Computational Harmonic Analysis, 2009, 26(3): 301-321.
- [30] Yang H G, Lin H B, Ding K. Sliding window denoising K-Singular Value Decomposition and its application on rolling bearing impact fault diagnosis [J]. Journal of Sound and Vibration, 2018, 421: 205-219.

Rolling bearing fault feature extraction in the compressed domain with deep convolutional measurement network

LIN Hui-bin, WANG Hong-chang, XI Ci-yang

(School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510641, China)

Abstract: Compressed sensing can effectively relieve the burden of data storage and transmission for mechanical condition monitoring. However, this method exists some problems such as low compression efficiency and slow signal reconstruction process in the application of fault diagnosis. In this paper, using the corresponding relationship between autoencoder and compressed sensing, a novel fault feature extraction method of the rolling bearing in the compressed domain based on the deep convolutional measurement network is proposed. For the problem that noise-free fault signal samples are difficult to obtain, a dataset construction method based on the fault mechanism is proposed. The model trained on this dataset is suitable for bearing signals under different working conditions. A deep convolutional denoising autoencoder (DCDAE) is constructed, in which the number of layers is determined by the required signal compression rate and the frequency of the hidden layer corresponds to that of the original signal. The fully trained encoding sub-network of DCDAE, named deep convolutional measurement network (DCMN), is used to compress the rolling bearing vibration signal instead of the traditional measurement matrix, and then the fault features are directly extracted in the compressed domain. The effectiveness of the proposed dataset construction method and the compressed domain feature extraction method are analyzed through the simulations. The rolling bearing experimental signals further verify that the deep convolutional measurement network trained by the proposed method has good generalization and can effectively extract fault features for fault diagnosis in the compressed domain with a compression ratio far lower than that of the traditional compressed sensing method.

Key words: fault diagnosis; rolling bearing; fault feature extraction; compressed sensing; deep convolutional measurement network

作者简介: 林慧斌(1975—),女,博士,副教授。电话:(020)87113220;E-mail:hblin@scut.edu.com。