DOI: 10.3901/JME.2020.07.096

基于深度一维残差卷积自编码网络的 齿轮箱故障诊断*

周兴康 余建波

(同济大学机械与能源工程学院 上海 201804)

摘要: 一维振动信号常常被用于齿轮箱的监测与故障诊断中,使得能及时地对齿轮箱维护以减少损失。因此,从一维振动信号中提取出关键故障特征决定了故障诊断模型的准确性与可靠性。典型的深度神经网络(deep neural network, DNN),如卷积神经网络已经在故障诊断中表现出良好的性能并得到了广泛的应用,但其监督式训练方式往往需要大量的标签数据而限制了其可应用性。因此,提出一种新的深度神经网络模型,一维残差卷积自编码器(1-dimension residual convolutional auto-encoder, 1DRCAE),成功应用于振动信号的无监督学习及故障特征提取,显著提高了齿轮箱的故障诊断率。首先,提出了一维卷积层与自编码器的有效集成方法,形成了深度一维卷积自编码器;其次,引入残差学习机制训练一维卷积自编码器,实现对一维振动信号有效地特征提取;最后,基于编码器提取的特征,使用少量标签数据进行分类微调实现齿轮箱故障模式识别。通过齿轮箱试验台采集的传感器数据进行实验验证表明,这种无监督学习方法具有良好的去噪能力和故障特征提取能力,其特征提取效果好于典型的深度神经网络,如深度置信网络(Deep belief network, DBN)和堆叠自编码网络(Stacked auto-encoders, SAE),同时故障诊断效果也优于一维卷积神经网络(1-dimension convolutional neural network, 1DCNN)。

关键词: 齿轮箱故障诊断; 特征学习; 深度学习; 卷积自编码器; 残差学习

中图分类号: TH132

Gearbox Fault Diagnosis Based on One-dimension Residual Convolutional Auto-encoder

ZHOU Xingkang YU Jianbo

(School of Mechanical Engineering, Tongji University, Shanghai 201804)

Abstract: One-dimension vibration signals are often used for gearbox monitoring and fault diagnosis to perform maintenance timely and then reduce losses. Thus, the accuracy and reliability of the fault diagnosis model are determined by the key fault features extracted from one-dimension vibration signals. Typical deep neural networks (DNNs), e.g., convolutional neural network, have performed well in fault diagnosis and been widely used in machine fault diagnosis. However, supervised learning of CNN often requires a large amount of labeled data thus limits its wide application. So a new DNN, one-dimension residual convolutional auto-encoder (1DRCAE), is proposed for unsupervised learning and feature extraction from vibration signals, which significantly improves diagnosis accuracy of the gearbox fault. First, the effective integration of the one-dimension convolutional layer and the auto-encoder is realized, and a deep one-dimension convolutional auto-encoder is constructed in the 1DRCAE network. Secondly, the residual learning mechanism is introduced to train the one-dimension residual convolutional auto-encoder to learn effective features from one-dimension vibration signals. Finally, based on the features extracted by 1DRCAE, a classifier is constructed using a small amount of labeled data for fine-tuning and then is used for the gearbox fault diagnosis. The vibration signals from a gearbox test rig are used to verify effectiveness of 1DRCAE for gearbox fault diagnosis. The experimental results show that this unsupervised learning of 1DRCAE has good denoising and feature extraction ability. It performs much better on feature extraction than the typical DNNs, e.g., deep belief network (DBN) and stacked auto-encoders (SAE). 1DRCAE also shows better performance on fault diagnosis than that of 1-D CNN (1-DCNN), etc.

Key words: gearbox fault diagnosis; feature learning; deep learning; convolutional auto-encoder; residual learning

* 国家自然科学基金(71777173)、中央高校基本科研业务费和上海科委创新科技行动计划(19511106303)资助项目。20190119收到初稿,20190627收到修改稿

0 前言

齿轮箱作为各种机械设备中的重要部件, 在运行过程中发生故障将会导致高昂的维护成本并带来安全问题。因此, 利用传感器收集的一维振动信号对齿轮箱进行实时的状态监控和故障诊断具有重要意义。随着机器学习的发展, 各种机器学习技术已被用于故障诊断, 例如人工神经网络^[1-3], 支持向量机^[1,4], 高斯混合模型^[5]等。这些机器学习方法仍需要人工进行特征选择^[6], 数据不能直接作为模型输入进行训练, 且模型诊断准确率受初始参数影响较大。

深度学习算法通过建立深度神经网络, 能够从海量数据中提取数据特征, 且提取到的数据特征随着网络层数的深入在变得更为抽象。作为监督式学习网络的代表, 深度卷积神经网络(Convolutional neural network, CNN)的训练需要大量带有标签的数据进行反向传播微调以获得更好性能。得益于 ImageNet^[7]的海量标签图像数据, AlexNet^[8]在 2012 年 ImageNet 图像识别挑战赛上大幅提高了识别准确率, 使得 CNN 受到研究人员广泛关注。同时深度学习也涌现出无监督式的学习算法, 如深度置信网络^[9](Deep belief network, DBN), 堆叠自编码网络^[10](Stacked auto-encoder, SAE), 生成式对抗网络^[11](Generative adverse network, GAN)等。这些无监督学习算法通过重构输入数据, 在特征学习阶段不需要借助标签数据。深度学习良好地提取特征的能力使其在图像分类^[12-14]、人脸检测^[15]、自然语言处理^[16]等领域均取得了很好的效果。

同时, 卷积神经网络的成功使得研究人员尝试将卷积核整合到其他神经网络中, 出现了如卷积自编码网络(Convolutional auto-encoder, CAE), 卷积深度信念网络(Convolutional deep belief network, CDBN)等混合学习的网络。MASCI 等^[17]采用计算单元为卷积核的自编码器训练得到的卷积核参数来初始 CNN 并成功应用于 MNIST^[18]识别上。采用卷积核作为自编码器计算神经元的网络在图像分类获得了很好的分类效果^[19-21]。TABAR 等^[22]训练一维 CNN 后再使用 SAE 网络替换全连接层。HUANG 等^[23]采用卷积核为自编码器的基本单元, 用于提取脑功能核磁共振数据中的特征。但这些方法不能直接用于一维的传感器信号故障诊断任务。

在故障诊断领域, 一些典型的深度神经网络得到了成功应用。JING 等^[24]对齿轮箱振动信号数据进行了时频域变化将得到的频域图像作为 CNN 的

输入来训练网络, 同时还讨论了不同网络结构对故障诊断率的影响。ZHANG 等^[25,26]直接使用轴承原始信号来训练一维卷积神经网络(1DCNN)进行故障诊断, 其采用宽卷积核来抑制高频噪声。GUO 等^[27]提出了层级卷积神经网络用于轴承故障诊断, 使用两个卷积网络分别用于学习信号特征和评估测试信号的故障模式。张西宁等^[28]使用卷积编码器训练得到的卷积核参数来初始化一维卷积神经网络, 运用于故障分类识别。SHAO 等^[29]使用受限玻尔兹曼机(Restricted boltzmann machine, RBM)学习数据特征, 并采用粒子群优化算法(Particle swarm optimization, PSO)优化深度置信网络结构应用于轴承故障诊断。SUN 等^[30]使用稀疏自动编码器提取感应电机特征, 利用学到的参数初始化分类网络进行故障诊断。孙文珺等^[31], LU 等^[32]对输入信号加入噪声训练编码器, 获得了鲁棒性更强的去噪编码器。但上述方法中, 采用卷积神经网络进行故障诊断的方法均属于监督式学习, 其泛化性和准确性依赖于使用大量的标签数据来训练网络避免过拟合。而无监督方法的网络连接方式多为全连接神经元, 相较于卷积计算, 其提取特征能力较差, 网络的训练方式较为繁琐。

本文提出了一种新的深度神经网络, 一维残差卷积自编码器(1-dimension residual convolutional auto-encoder, 1DRCAE), 并提出了基于 1DRCAE 特征学习的齿轮箱故障诊断模型, 主要贡献如下。

(1) 1DRCAE 采用一维卷积核为基本计算单元, 可直接将原始传感器信号作为输入并应用于故障诊断中, 相较于传统机器学习, 网络无需标签数据即可学习并提取数据的关键特征。

(2) 相较于 SAE 等无监督算法, 一维卷积核学习到的数据特征比全连接神经元所学习到的特征更为抽象, 同时网络中引入的残差学习机制改善了训练过程, 引入残差学习后的端到端的训练方式提高网络特征提取的能力。

(3) 相较于 CNN 等监督式算法, 1DRCAE 通过学习大量无标签数据中的特征, 仅需要少量的标签数据来训练分类器并进行故障模式识别, 对于标签数据量的需求大大降低。

1 理论背景

1.1 一维卷积神经网络

卷积神经网络基于哺乳动物的视觉皮层研究, 是一种具有层级结构的深度神经网络。一维卷积神经网络主要由一维卷积层、一维池化层、全连接层和分类器构成, 如图 1 所示。

图1 一维卷积神经网络

卷积层: 对一维输入信号 $X \in R^L$, 一维卷积神经网络的卷积层使用 K 个宽度为 w 的卷积核 $\omega_i \in R^w (i=1,2,\dots,K)$ 对其进行一维卷积, 其输出

$$out_put_i = f\left(\sum X \odot \omega_i + b_i\right) \quad i=1,\dots,K \quad (1)$$

式中, $\odot$ 代表卷积核与输入的卷积运算, f 为非线性激活函数, b_k 为每个通道的偏置, K 为输出后的通道数。

池化层: 通常选用最大池化方法, 对数据 $T \in R^{K \times L}$, 池化后输出

$$P_i(n) = \max_{0 \leq n \leq \frac{L}{S}} \{T_i(nW, (n+1)W)\} \quad i=1,\dots,K \quad (2)$$

其中 $T_i \in R^L$ 表示第 i 个特征张量, W 是池化窗口的大小, S 为步长。

全连接层: 参数为权重 ω 和偏差 b , f 为非线性激活函数, 对输入 $P \in R^m$, 全连接层的输出为

$$\delta = f(\omega P + b) \quad (3)$$

分类器层: Softmax 得到输入数据的标签分布

$$Q(\delta^i) = \frac{e^{\delta^i}}{\sum_k e^{\delta^k}} \quad (4)$$

1.2 自编码网络

自编码网络是一种无监督学习网络, 网络分为编码器和解码器两个部分, 有全连接神经元构成, 如图2所示。

编码阶段, 给定输入数据 $X \in R^L$, 编码器将输入变换为特征矢量 h

$$h = f(\omega X + b) \quad (5)$$

解码阶段, 通过解码器重构输入, 得到

$$\hat{X} = f(\omega' h + b') \quad (6)$$

式中, ω 和 b 为网络的权重和偏差, f 为非线性激活函数, 网络的损失函数为

$$Loss = \|X - \hat{X}\|^2 \quad (7)$$

图2 自编码器结构

1.3 残差学习

HE 等^[33]提出了残差卷积神经网络(Residual neural network, ResNet)取得了当年 ImageNet 竞赛最好成绩。ResNet 的核心思想是引入残差模块来学习残差。残差学习解决了深度卷积神经网络随着层数的增加网络性能反而下降的现象。残差模块的定义为

$$y = F(x, \{W_i\}) + x \quad (8)$$

式中, x 和 y 为模块的输入和输出, F 表示将要学习的残差映射, W_i 表示模块参数。

如图3所示, 残差学习的不再是对该层输入的潜在映射 $H(X)$, 而是残差映射 $F(X)=H(X)-X$ 。图中将输入 X 恒等映射到模块输出的连接称为跳跃连接^[34], 这种连接能够更好地传播梯度, 以解决深层卷积网络的训练问题同时提高网络性能, 如 Inception-V4^[35], Res-Next^[36]等都使用了残差学习来提高网络的性能。

图3 残差模块示意图

2 一维残差卷积自编码器

2.1 网络结构

1DCNN 的训练需要用到大量的标签数据, 获得这些标签数据卷积往往是需要耗费很多的时间和费用。1DRCAE 基于自编码器无监督学习的思想, 通过卷积和反卷积等运算来提取数据潜在特征, 在微调阶段使用少量标签数据训练分类器, 网络结构如图4所示。

图 4 一维残差卷积自编码器

1DRCAE 的无监督学习分为编码和解码两个阶段。图 5 为网络采用的基本计算单元。编码阶段采用卷积和池化对数据进行编码，解码阶段采用了反卷积和上采样对数据进行还原。网络采用一维卷积核作为基本运算单元，可以直接使用一维原始信号；引入残差学习机制改善网络的梯度传导，从而控制重构误差，提高网络特征提取能力。

图 5 网络使用的运算方式

(1) 编码部分由一维卷积层、一维池化层和瓶颈层构成。一维卷积层如图 5a 所示，对第 i 个一维卷积核，其输出为第 i 个通道的特征

$$C_i = \text{ReLU}\left(\sum X \cdot \omega_i + b_i\right) \quad (9)$$

其中 ReLU 为整流线性单元(Rectified Linear Unit)

$$\text{ReLU}(x) = \max(0, x) \quad (10)$$

一维池化层如图 5c 所示，对数据进行降维，对

长度为 1 的第 i 个通道的特征，池化后的输出为

$$P_i(n) = \max_{0 \leq n \leq \frac{L}{S}} \{T_i(nW, (n+1)W)\} \quad (11)$$

其中池化窗口大小 $W=2$ ，步幅 $S=2$ ，池化后每个通道特征的长度变为原来的一半。

瓶颈层为卷积核大小和步幅均为 1，卷积核数量较少的卷积层，这种 1×1 的卷积连接可以进行通道上的降维，且瓶颈层的参数可以学习，瓶颈层可以减少全连接层的神经元，进而降低分类器大小。

(2) 解码部分主要由反卷积层和上采样层组成，与编码部分的运算相逆。

反卷积层如图 5b 所示，由公式(9)可反向推导出反卷积层第 i 个通道的输出为

$$D_i = \text{ReLU}\left(\sum X \otimes \omega_i + b_i\right) \quad (12)$$

式中， ω_i 代表反卷积核， $\otimes$ 表示反卷积核在输入特征 X 上滑动进行反卷积运算。

上采样层如图 5d 所示，其输出是对上一层输入数据的放大，以恢复池化之后减半的数据维度。对输入的第 i 个特征，其输出为

$$U_k^i = \begin{cases} 0 & k \neq j_k \\ X^i & k = j_k \end{cases} \quad k \in [t, 2t] = 1, 2, \dots, l \quad (13)$$

式中， l 为特征长度， j_k 为池化时记录的最大值的位置。

(3) 残差模块通过跳跃连接传递数据特征，对于图 4 中的残差模块(Residual Block)，上采样层 2(Upsample2)的输入

$$X_{u2} = D^1(Xd_1) + P^1(Xp_1) = D^1(U^1(X_{u1})) + P^1(C^1(X)) \quad (14)$$

式中, D^1 为反卷积层 1, P^1 为池化层 1, U^1 为上采样层 1, C^1 为卷积层 1, X 为原始输入数据, $X_{d,p,u}$ 为其所在层的输入。

从上式可以看到, X_{u2} 不仅由上一层的反卷积和上采样决定, 还保留了原始数据 X 经过卷积和池化后的特征, 使得解码器可以更好地重构数据。

对于上采样层 2, 其输出为

$$y_{u2} = U^2(X_{u2}) \quad (15)$$

且 y_{u2} 作为反卷积层 2 的输入, 得到输出

$$\bar{X} = D^2(X_{d2}) = D^2(y_{u2}) \quad (16)$$

从式(14)~(16)中可以看到, 1DCNN 网络的输入和输出拥有了一条更短的连接, 当使用梯度下降法求解整个网络参数时, 这种连接可以避免优化中的梯度问题。对于靠近输入数据的网络第一层, 卷积层 1 获得的梯度也由两部分构成, 其改善编码器的特征提取的训练。

2.2 参数优化

1DRCAE 通过重构输入信号来学习特征, 其损失函数为

$$Loss = \text{avg}|X - \bar{X}|^2 \quad (17)$$

式中, X 和 $\bar{X}$ 分别为编码器的输入和输出。

为了最小化输入和输出的均方误差解码阶段, 采用梯度下降法来优化网络参数。

对于反卷积层, 其参数为反卷积核, 梯度计算如下

$$\frac{\partial L}{\partial K_i(j)} = \frac{\partial L}{\partial D} \frac{\partial D}{\partial K_i(j)} \quad (18)$$

$$\frac{\partial D}{\partial K_i(j)} = \begin{cases} 0, & i \neq j + km \\ x_{k+1}, & i = j + km \end{cases} \quad k = 0, 1, \dots, t \quad (19)$$

式中, $K(j)$ 为第 j 个反卷积核, x 、 D 分别为反卷积的输入和输出, t 为 x 长度, m 为反卷积步长。

对于上采样层, 由于其不含参数仅对数据进行扩增, 因此梯度传递如下

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= \frac{\partial L}{\partial U_i} \frac{\partial U_i}{\partial x_i} = \frac{\partial L}{\partial U_i} \cdot 1 = \\ &\frac{\partial L}{\partial U_i} \quad i \in [tk, (t+1)k] \end{aligned} \quad (20)$$

式中, x_i 表示该层第 i 个输入, U_i 表示第 t 个输出, k 表示上采样的步长。

同理, 对于池化层, 梯度传递如下

$$\frac{\partial L}{\partial x_i} = \frac{\partial L}{\partial P_i} \frac{\partial P_i}{\partial x_i} = \begin{cases} \frac{\partial L}{\partial P_i} x_i = \max(x) \\ 0 & x_i = \max(x) \end{cases} \quad (21)$$

式中, x , P 分别为池化层的输入和输出。

对于卷积层, 计算卷积核的梯度

$$\begin{aligned} \frac{\partial L}{\partial K_k(j)} &= \frac{\partial L}{\partial C} \frac{\partial C}{\partial K_k(j)} = \\ &\sum (x_i) \cdot \text{flip} \left(\frac{\partial L}{\partial C} \right) \end{aligned} \quad (22)$$

式中, x , C 表示卷积层的输入和输出, k 表示第 k 个卷积核, $\odot$ 表示卷积计算, flip 表示矩阵的翻转。

原位于反向传播过程尾端的卷积层 1, 引入残差模块后, 误差对该层输出 C^1 的梯度变为

$$\begin{aligned} \nabla C &= \frac{\partial L}{\partial C^1} = \frac{\partial L}{\partial P^1} \cdot \frac{\partial P^1}{\partial C^1} = \\ &\left(\frac{\partial L}{\partial C^2} \cdot \frac{\partial C^2}{\partial P^1} + \frac{\partial L}{\partial U^2} \cdot \frac{\partial U^2}{\partial X_u} \right) \frac{\partial P^1}{\partial C^1} \end{aligned} \quad (23)$$

式中, C , U , P , D 分别为四种不同层的输出, 上标代表图 4 中的 1 或 2 层, X_u 为上采样层输入。

由误差对 C^2 的逐层传递的梯度

$$\frac{\partial L}{\partial C^2} = \frac{\partial L}{\partial D^2} \frac{\partial D^2}{\partial U^2} \cdots \frac{\partial P^2}{\partial C^2} = \nabla H \quad (24)$$

得到卷积层 1 的梯度为

$$\begin{aligned} \nabla C &= \nabla H \cdot \frac{\partial C^2}{\partial P^1} \cdot \frac{\partial P^1}{\partial C^1} + \frac{\partial L}{\partial U^2} \cdot \frac{\partial U^2}{\partial X_u} \cdot \frac{\partial P^1}{\partial C^1} = \\ &\nabla H' \cdot \frac{\partial L}{\partial U^2} \cdot \frac{\partial U^2}{\partial X_u} \cdot \frac{\partial P^1}{\partial C^1} \end{aligned} \quad (25)$$

从上式可以看出, 传递到卷积层 1 的梯度除了逐层传递的梯度之外, 还包含了损失经过跳跃连接传递的梯度。这种跳跃连接有效地改善了深层网络传播的梯度问题。作为特征提取器的第一层, 卷积层 1 学到的参数决定了整个特征提取的质量, 因此这种梯度改善有效地提高了特征提取器的性能。

在微调阶段, 将编码后的特征输入全连接层和分类器, 输出预测标签数据。得到故障分类阶段损失函数为预测标签与真实标签分布的交叉熵

$$Loss = - \sum p(x) \log q(x) \quad (26)$$

式中, q 为真实标签分布, p 为由公式(4)Softmax 函数得到预测标签分布。为了优化该损失函数, 首先计算损失函数对 Softmax 函数的梯度

$$\frac{\partial L}{\partial z^{l+1(j)}} = \sum_{k=1}^m p_k^j q_k^j - p_k^j \quad (27)$$

最后再计算全连接层权重 ω 和偏差 b 的梯度并更新

$$\frac{\partial L}{\partial \omega_{ij}^l} = \frac{\partial L}{\partial z^{l+1(j)}} \cdot \frac{\partial z^{l+1(j)}}{\partial \omega_{ij}^l} = \frac{\partial L}{\partial z^{l+1(j)}} \cdot x^{l(i)} \quad (28)$$

$$\frac{\partial L}{\partial b_j^l} = \frac{\partial L}{\partial z^{l+1(j)}} \cdot \frac{\partial z^{l+1(j)}}{\partial b_j^l} \quad (29)$$

1DRCAE 的训练和测试过程如算法 1 所示。

算法 1: 1DRCAE 的训练和测试过程。

输入: 训练集 X , 训练集标签 y , 测试集 $\tilde{X}$

步骤 1: 无监督训练

1) 设定卷积层、池化层、反卷积层、上采样层以及学习率等超参数

2) 随机初始化编码器网络所有权重和偏差

3) For 训练次数 N

4) 输入 X_i

5) 计算卷积、池化输出特征 f_i

6) 计算上采样、反卷积输出 X'_i

7) 计算误差 $\text{avg}|X_i - X'_i|^2$

8) 求解反卷积参数梯度 t_i

9) 更新反卷积参数 ω^T, b^T

10) 求解卷积梯度参数 d_i

11) 更新卷积参数 ω, b

12) End For

步骤 2: 分类微调

1) 初始化全连接层权重和偏差

2) For 训练次数 M

3) 输入 X_i 、标签 y_i

4) 卷积、池化输出特征 f_i

5) f_i 输入全连接和分类器

6) 输出预测标签 y'_i

7) 计算损失 $-\sum_x y'_i \log y_i$

8) 求解全连接梯度

9) 更新全连接权重和偏差

10) End For

输出: $\tilde{X}$ 的预测类标签

1DCNN 网络训练使用随机梯度下降方法, 使用 Adam 优化算法进行优化, 如算法 2 所示。

算法 2: Adam 优化算法。

默认参数: $\alpha=0.001, \beta_1=0.9, \beta_2=0.999, \varepsilon=10^{-8}$

$m_0 \leftarrow 0$ (初始化第 1 个动量参数)

$v_0 \leftarrow 0$ (初始化第 2 个动量参数)

$t \leftarrow 0$ (初始化时间步)

while 网络收敛不满足要求(θ_i 表示网络所有权重和偏置)

$t \leftarrow t+1$

$g_t \leftarrow \nabla \theta f_t(\theta_{t-1})$ (计算损失对网络参数的梯度)

$m_t \leftarrow \beta_1 \cdot m_{t-1} + (1-\beta_1) \cdot g_t$ (更新带有偏差的第 1 动量估计)

$v_t \leftarrow \beta_2 \cdot v_{t-1} + (1-\beta_2) \cdot g_t^2$ (更新带有偏差的

第 2 动量估计)

$\hat{m}_t \leftarrow m_t / (1-\beta_1^t)$ (计算修正后的第 1 动量估计)

$\hat{v}_t \leftarrow v_t / (1-\beta_2^t)$ (计算修正后的第 2 动量估计)

$\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)$ (更新网络参数)

end while

输出: 网络模型参数

2.3 基于 1DRCAE 的故障诊断流程

基于 1DRCAE 的故障诊断流程如图 6 所示, 分为离线和在线两个部分。离线训练时, 使用训练数据训练特征提取器和分类器; 在线检测时, 网络对输入数据进行分类预测并输出标签。

图 6 基于 1DRCAE 的齿轮箱故障诊断流程

3 实验与结果分析

为了验证 1DRCAE 的有效性, 使用行星齿轮箱实验台采集到的振动信号数据进行故障诊断。实验台如图 7 所示, 包含行星齿轮箱(含一个 75 齿太阳齿轮和 25 齿行星齿轮)、驱动电机和磁力制动器。诊断信号由固定在变速箱壳体上的加速度传感器收集, 采样频率为 30 kHz。

图 7 行星齿轮箱实验台

行星齿轮箱的健康状况为正常、齿牙凹陷,不同程度的齿牙磨损,齿牙裂纹和 3 种模式的轴承缺陷。带有不同缺陷的齿轮和轴承如图 8 所示。

图 8 缺陷齿轮和轴承

实验采集到的振动信号每类有 190 多万个点,我们以 2 048 个点作为一个输入样本,得到了含有 6 336 个样本的训练集和 3 795 个样本的测试集。11 种不同健康状况的时域信号如图 9 所示。对每种故障模式数据的具体描述如表 1 所示。

图 9 十一种齿轮箱状态的时域信号

表 1 数据描述

模式类别	齿轮箱状况	转速/(r/min)	训练集	测试集
C1	正常	294,588, 882	576	345
C2	齿牙磨损 1	294,588, 882	576	345

(续)

模式类别	齿轮箱状况	转速/(r/min)	训练集	测试集
C3	齿牙磨损 2	294,588, 882	576	345
C4	齿牙磨损 3	294,588, 882	576	345
C5	齿牙磨损 4	294,588, 882	576	345
C6	齿牙磨损 5	294,588, 882	576	345
C7	齿牙点蚀	294,588, 882	576	345
C8	齿牙断裂	294,588, 882	576	345
C9	轴承内圈缺陷	294,588, 882	576	345
C10	球滚子缺陷	294,588, 882	576	345
C11	保持架缺陷	294,588, 882	576	345

3.1 实验部分

1DRCAE 的网络参数设定如表 2 所示,在网络的第 1 层卷积核中,选用了较宽的卷积核来过滤噪声,设置步幅为 2 来进行数据长度上的降维。在池化层的参数设置上,设定池化窗口大小及步幅均为 2,除了保持特征的局部不变以外,还降低了数据维度。同时增加了瓶颈层,其作用为降低卷积池化输出后的通道维度,因此卷积核大小和步幅均为 1。同时保持整个网络的对称结构,解码器的参数设置与编码器类似。

表 2 网络参数

1D-RCAE	网络层	参数
结构参数	卷积层 C1	宽度=64, 步幅=2
	池化层 P1	宽度=2, 步幅=2
	卷积层 C2	宽度=4, 步幅=2
	池化层 P2	宽度=2, 步幅=2
	瓶颈层 1&2	宽度=1, 步幅=1
	上采样层 U1	宽度=2, 步幅=2
	反卷积层 D1	宽度=4, 步幅=2
	上采样层 U2	宽度=2, 步幅=2
训练参数	反卷积层 D2	宽度=64, 步幅=2
	学习率	0.000 1
	批次大小	100

训练阶段的损失函数下降如图 10 所示,无监督训练后,使用 1DRCAE 对几类数据进行了重构,重构输出和原始信号波形图对比如图 11 所示,可以看到经过编码器重构后的信号其关键特征变得显著,同时噪声得到有效消除,在下节可视化分析中可以看到这种去噪效果来源于第一层充当着滤波器作用的宽卷积核的滤波作用。这个结果充分说明 1DRCAE 具有滤除原始振动信号中噪音,从而提取关键故障特征的能力。

通过无监督学习,比较原始信号输入与 1DRCAE 输出可以看到网络对原始信号的重构较

好，证明网络中间层提取到的特征具有较强的代表性，网络训练时的故障诊断正确率如图 12 所示。

(a) 均方误差下降

(b) 分类交叉熵下降

图 10 训练过程损失下降图

图 11 几类数据的原始输入和 1DRCAE 重构输出比较

可以看到，当训练次数 2 000 次时，分类正确率就已经超过 80，随着训练次数的不断增加，

1DRCAE 的最终分类正确率在测试集上达到了 96.44%。1DRCAE 故障诊断的混淆矩阵如图 13 所示。1DRCAE 在每个类上均获得非常好的识别结果。

图 12 模型训练准确率

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1	91.8	0	0	0	0	0	0	0	0	0	0
C2	0	97.8	0.8	0	0	0	0	0	0	0	0
C3	0	0	98.4	0.7	0	0	0.8	0	0	0	0
C4	0	0	0	98	0	0	0.8	0	0	0	0
C5	0	0.7	0	1.3	96.4	7.3	0	0	0	0	0
C6	0	0	0	0	3.6	92.7	0	0	0	0	0
C7	1.4	0	0.8	0	0	0	98.4	0	1.8	0	0
C8	0	0.7	0	0	0	0	0	99.3	0	0	0.8
C9	6.8	0.7	0	0	0	0	0	0.7	97.3	0	4.2
C10	0	0	0	0	0	0	0	0	0	100	0
C11	0	0	0	0	0	0	0	0	0.9	0	95

图 13 识别率混淆矩阵(%)

与典型的 DNN 模型 SDAE 与 DBN 的五折交叉对比结果如图 14 所示，可以看到由于该方法具有更好的特征提取能力，分类准确率也更高。

SDAE	94.6	97.2	96.7	90.6	83.6	80.8	96.3	92.3	80.8	85.8	78	88.9
	97	97	99.7	85.5	79.1	78.8	98.5	92.3	76.4	83.8	85	88.3
	96.9	94.9	93.4	89.2	80.4	79.1	98.1	97	82	80.3	82.3	88.4
	97.3	96.2	97.4	89.4	72.3	80.8	90.3	95.5	76.3	85.2	90.5	88.3
	96	96.8	97.7	88.7	71.5	82.6	98.4	95.8	67.4	87.8	84.5	88.2
DBN												88.42
	96	99	95.2	96.8	83.2	78.7	97.5	91.5	85.3	88.4	85.2	90.7
	98.7	99.2	96.9	98.2	73	84.6	95.9	96.2	81.1	90.1	85.9	90.8
	96.7	97.8	96.1	98.5	71.1	88.1	94.7	92.4	70.8	91.8	86.4	89.2
	99.2	98.7	94.1	98.7	78.2	80.7	95.3	96	78.7	89.3	86.1	90.5
1DRCAE	96.7	97.6	95.4	97.5	76.8	82.7	96.6	96.3	72.3	87.8	84.2	89.7
												90.18
	86.1	99	99.2	99.7	97.8	96.3	99.7	99.1	95.7	98.5	95.6	96.7
	91.9	99.5	98.4	99.5	93.1	93.6	96.2	93.8	96.4	99.8	97.7	96.4
	88.3	97.8	98.3	99.7	97.5	96.6	98.4	95.8	95.3	99.5	94	96.4
	86.9	99.2	98.3	97.1	96.8	97.7	98.8	97.2	91.9	100	93.2	96.2
	90.6	98.7	97.8	99	91.4	94.7	99.7	97.7	97.5	98.7	95.1	96.5
												96.44

图 14 典型的神经网络对比(%)

为了进一步验证该方法的有效性和先进性，在齿轮箱数据上使用故障诊断领域常用的监督式学习算法作为对比。其中：一维 CNN 的层数为 6(含有 2 层卷积层、2 层池化层和 2 层全连接层)；

全连接网络(BPN)的层数为 3(神经元数分别为 2 048、1 024、11); 随机森林的最大深度设置为 50, 包含 800 棵树。

各个方法五折交叉最终对比结果如表 3 所示。

从结果可以看出, 1DRCAE 引入的残差学习和卷积运算使得提取特征的能力更好, 故障诊断效果更佳, 其识别率显著地好于其他 DNN 和典型识别器。

3.2 可视化分析

本节对网络进行可视化分析从而了解 1DRCAE 如何进行关键特征提取和分类。

图 15 第一层卷积核波形图

图 16 第 7 类故障的特征输出

3.2.1 卷积可视化

在 1DRCAE 编码器网络采用了宽卷积核作用于输入数据上, 实验表明, 网络对输入的重构结果相比于原始输入其关键特征更为明显且噪声也显著降低。将卷积核参数可视化得到的第一层的 16 个卷积核, 如图 15 所示, 卷积核表现为不同波形的滤波器, 这种滤波具有一定的降噪作用。且通过不断的反向传播训练, 卷积核可以学习到振动信号的关键故障特征。

以第 7 类故障数据作为分析对象输入编码器, 经过卷积核池化层编码后的特征可视化如图 16 所示, 可以看到, 仅有部分卷积核对该类

故障数据响应(如图 1-6),一部分卷积核的输出为零(如图 1-7),这得益于线性整流单元(ReLU)的选用。

表 3 各类典型识别器的对比结果(%)

	1DRCAE	1DCNN	DBN	SDAE	BPN	RF
C1	88.76	85.20	97.46	96.36	97.80	83.50
C2	98.84	100.00	98.46	96.42	98.50	94.50
C3	98.40	89.10	95.54	96.98	93.50	86.40
C4	99.00	94.70	97.94	88.68	97.00	96.30
C5	95.32	94.70	76.46	77.38	68.70	84.00
C6	95.78	94.90	82.96	80.42	84.80	72.40
C7	98.56	87.00	96.00	96.32	96.70	97.50
C8	96.72	89.80	94.48	94.58	86.70	97.80
C9	95.36	94.60	77.64	76.58	62.50	78.70
C10	99.30	98.60	89.48	84.58	82.50	83.60
C11	95.12	99.10	85.56	84.06	75.60	58.80
AVG	96.44	93.40	90.18	88.42	86.30	85.50

3.2.2 特征可视化

为了验证 1DRCAE 逐层提取特征能力,采用 t-SNE 对提取特征进行降维并可视化呈现。如图 17 所示,图 17a~17d 分别代表原始输入数据,第一层卷积池化后输出,第二层卷积池化后输出和全连接层输出。可以看出经过卷积编码后,相同标签的特征数据在二维空间上具有了较好相似性。而经过微调阶段的特征选择后,不同标签的数据得到了很好的分类效果。

图 17 t-SNE 降维可视化

4 参数敏感性分析

首先分析残差学习机制对于自编码器的影响,整个自编码器的前 500 步均方误差下降如图 18 所示,可以看到,引入残差后,编码器对数据的重构误差下降速度明显加快,且迭代结束后的误差小于未引入残差时的编码器误差。这表明残差学习机制对于提取特征并用于重构数据起到了重要作用。

图 18 引入残差学习机制的自编码器损失下降

考虑到对于一维振动信号,1DRCAE 的卷积核具有滤波器的作用,因此分析卷积核宽度即滤波器长度对于特征提取和故障诊断的影响。对比结果分别如图 19 与图 20 所示。为了比较误差的相对下降速率,图 19 中的损失为相对数据的长度(即除以 2048)后的结果。结果表明较宽的卷积核更有利于提取一维信号的关键特征,但同时过宽的卷积核会导致参数冗余而难以训练,影响网络的特征提取。

图 19 采用不同卷积核的损失下降

图 20 采用不同卷积核的故障识别准确率

不同的卷积核宽度也影响振动信号特征提取的结果,如表 4 所示。不同宽度卷积核的 1DRCAE 识别正确性不同,第一层卷积编码层采用较宽的卷积核可以很好地学习振动信号特征,从而保证了编码器和分类器的性能。同时,当卷积核的宽度增加到一定宽度时,特征的提取效果不再有效增加,并不能带来更好的分类效果,同时过大的卷积核也会浪费计算资源,所以本文采用宽度为 64 的卷积核作为第一层卷积层的计算单元较为合理。

表 4 卷积核宽度对故障识别率的影响 (%)

卷积核宽度	4	16	64	128
准确率	92.6	93.7	96.44	96.0

5 结论

本文提出了一个新的 DNN 模型,即 1DRCAE 其有效地融合了一维卷积核的残差卷积编码器,可以自适应地无监督学习传感器信号的抽象特征,从而有效地进行齿轮箱故障诊断,主要结论如下。

(1) 采用一维卷积核的自编码器可以有效地取出振动信号的噪音,并有效地提取振动信号的关键特征,从而有效地提高了识别器的性能;

(2) 引入的残差学习机制改善了自编码器的训练过程,使得可以端到端地训练一个深度神经网络。同时残差学习还加快了自编码器梯度的反向传播,减少了自编码网络的重构误差,提高了网络提取特征的能力;

(3) 1DRCAE 相比于其他算法不需要大量标签数据,更具有实用价值。良好的特征学习能力使得其可广泛用于齿轮箱等故障诊断领域。

参 考 文 献

- [1] SAMANTA B. Gear fault detection using artificial neural networks and support vector machines with genetic algorithms[J]. Mechanical Systems and Signal Processing, 2004, 18(3): 625-644.
- [2] ABU-MAHFOUZ I A. A comparative study of three artificial neural networks for the detection and classification of gear faults[J]. International Journal of General Systems, 2005, 34(3): 261-277.
- [3] YANG Shulian, LI Wenhui, WANG Canlin. The intelligent fault diagnosis of wind turbine gearbox based on artificial neural network[C]// International Conference on Condition Monitoring & Diagnosis. IEEE, Beijing, China, 2008: 1327-1330.
- [4] 袁胜发, 褚福磊, 何永勇. 基于网格支持矢量机的涡轮泵多故障诊断[J]. 机械工程学报, 2007, 43(4): 152-158. YUAN Shengfa, CHU Fulei, HE Yongyong. Multi-fault diagnosis for turbo-pump based on mesh support vector machines[J]. Chinese Journal of Mechanical Engineering, 2007, 43(4): 152-158.
- [5] WANG Guofeng, LI Yubo, LUO Zhigao. Fault classification of rolling bearing based on reconstructed phase space and Gaussian mixture model[J]. Journal of Sound and Vibration, 2009, 323(3): 1077-1089.
- [6] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21): 49-56. LEI Yaguo, JIA Feng, ZHOU Xin, et al. A deep learning based method for machinery health monitoring with big data[J]. Chinese Journal of Mechanical Engineering, 2015, 51(21): 49-56.
- [7] DENG Jia, DONG Wei, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]// Proceedings of the IEEE Computer Vision and Pattern Recognition, 2009: 248-255.
- [8] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Advances in Neural Information Processing Systems, 2012: 1106-1114.
- [9] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [10] RANZATO M, BOUREAU Y L, LECUN Y. Sparse feature learning for deep belief networks[C]// Advances in

- Neural Information Processing Systems, 2008: 1185-1192.
- [11] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// Advances in Neural Information Processing Systems, 2014: 2672-2680.
- [12] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2016: 2818-2826.
- [13] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// International Conference on Learning Representations (ICLR), 2015.
- [14] HUANG Gao, LIU Zhuang, MAATEN L V D, et al. Densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700-4708.
- [15] ZHANG Kaipeng, ZHANG Zhanpeng, LI Zhifeng, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [16] TAI Kaisheng, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long short-term memory networks[J]. Computer Science, 2015, 5(1): 36.
- [17] MASCI J, MEIER U, DAN Cireşan, et al. Stacked convolutional auto-encoders for hierarchical feature extraction[C]// International Conference on Artificial Neural Networks, Berlin, German, 2011: 52-59.
- [18] LECUN Y L, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [19] HAN Xiaobing, ZHONG Yanfei, ZHAO Bei, et al. Unsupervised hierarchical convolutional sparse auto-encoder for high spatial resolution imagery scene classification[C]// International Conference on Natural Computation, IEEE, 2016: 42-46.
- [20] CHEN Min, SHI Xiaobo, ZHANG Yin, et al. Deep features learning for medical image analysis with convolutional autoencoder neural network[J]. IEEE Transactions on Big Data, 2017: 1-1.
- [21] LUO Wei, LI Jun, YANG Jian, et al. Convolutional sparse autoencoders for image classification[J]. IEEE Trans Neural Netw Learn Syst, 2018, 29(7): 3289-3294.
- [22] TABAR Y R, HALICI U. A novel deep learning approach for classification of EEG motor imagery signals[J]. Journal of Neural Engineering, 2017, 14(1): 016003.
- [23] HUANG Heng, HU Xintao, ZHAO Yu, et al. Modeling task fMRI data via deep convolutional autoencoder[J]. IEEE Transactions on Medical Imaging, 2017, 37(7): 1551-1561.
- [24] JING Luyang, ZHAO Ming, LI Pin, et al. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox[J]. Measurement: Journal of the International Measurement Confederation, 2017, 111: 1-10.
- [25] ZHANG Wei, PENG Gaoliang, LI Chuanhao, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals[J]. Sensors, 2017, 17(2): 425.
- [26] ZHANG Wei, LI Chuanhao, PENG Gaoliang, 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: 439-453.
- [27] GUO Xiaojie, CHEN Liang, SHEN Changqing. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis[J]. Measurement, 2016, 93: 490-502.
- [28] 张西宁, 向宙, 唐春华. 一种深度卷积自编码网络及其在滚动轴承故障诊断中的应用[J]. 西安交通大学学报, 2018, 52(7): 1-8.
- ZHANG Xining, XIANG Zhou, TANG Chunhua. A deep convolutional auto-encoding neural network and its application in bearing fault diagnosis[J]. Journal of Xi'an Jiaotong University, 2018, 52(7): 1-8.
- [29] SHAO Haidong, JIANG Hongkai, ZHANG Xun, et al. Rolling bearing fault diagnosis using an optimization deep belief network[J]. Measurement Science and Technology, 2015, 26(11): 115002.
- [30] SUN Wenjun, SHAO Siyu, ZHAO Rui, et al. A sparse auto-encoder-based deep neural network approach for induction motor faults classification[J]. Measurement, 2016, 89: 171-178.
- [31] 孙文珺, 邵思羽, 严如强. 基于稀疏自动编码深度神经网络的感应电动机故障诊断[J]. 机械工程学报, 2016, 52(9): 65-71.
- SUN Wenjun, SHAO Siyu, YAN Ruqiang. Induction motor fault diagnosis based on deep neural network sparse auto-encoder[J]. Journal of Mechanical Engineering, 2016, 52(9): 65-71.
- [32] LU Chen, WANG Zhenya, QIN Weili, et al. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification[J].

- Signal Processing, 2017, 130: 377-388.
- [33] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [34] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Identity mappings in deep residual networks[C]// European Conference on Computer Vision(ECCV), 2016: 630-645.
- [35] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, Inception-ResNet and the impact of residual connections on learning[C]// The Thirty-First AAAI Conference on Artificial Intelligence(AAAI), 2017: 4278-4284.
- [36] XIE Saining, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5987-5995.
-
- 作者简介: 周兴康, 男, 1996 年出生。主要研究方向为深度学习和故障诊断。
E-mail: xkzhou@tongji.edu.cn
- 余建波 (通信作者), 男, 1978 年出生, 博士, 教授, 博士研究生导师。主要研究方向为故障诊断。
E-mail: jbyu@tongji.edu.cn