

控制与决策

Control and Decision

改进集成深层自编码器在轴承故障诊断中的应用

陈志刚, 杜小磊, 王衍学, 张楠

引用本文:

陈志刚, 杜小磊, 王衍学, 等. 改进集成深层自编码器在轴承故障诊断中的应用[J]. 控制与决策, 2021, 36(1): 135–142.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0270>

您可能感兴趣的其他文章

Articles you may be interested in

[结合注意力机制的循环神经网络复述识别模型](#)

Recurrent neural networks based paraphrase identification model combined with attention mechanism

控制与决策. 2021, 36(1): 152–158 <https://doi.org/10.13195/j.kzyjc.2019.0638>

[基于改进堆叠自动编码器的循环冷却水系统工艺介质温度预测控制方法](#)

Predictive control method of process medium temperature in circulating cooling water system based on improved stacked auto encoders

控制与决策. 2020, 35(12): 2835–2844 <https://doi.org/10.13195/j.kzyjc.2019.0694>

[事件触发机制下分布时滞网络化控制系统 \$H_\infty\$ 故障检测](#)

Event-triggered H_∞ fault detection for networked control systems with distributed delays

控制与决策. 2020, 35(12): 3059–3065 <https://doi.org/10.13195/j.kzyjc.2019.0456>

[参数未知的离散系统Q-学习优化状态估计与控制](#)

Q-learning optimal state estimation and control for discrete systems with unknown parameters

控制与决策. 2020, 35(12): 2889–2897 <https://doi.org/10.13195/j.kzyjc.2019.0180>

[基于不变网络模型和故障注入的分布式信息系统故障溯源方法](#)

Fault source location algorithm for distributed information system based on invariant network and fault injection

控制与决策. 2020, 35(11): 2723–2732 <https://doi.org/10.13195/j.kzyjc.2019.0214>

改进集成深层自编码器在轴承故障诊断中的应用

陈志刚^{1,2†}, 杜小磊^{1,2}, 王衍学^{1,3}, 张楠¹

(1. 北京建筑大学机电与车辆工程学院, 北京 100044; 2. 北京市建筑安全监测工程技术研究中心, 北京 100044; 3. 北京建筑大学城市轨道交通车辆服役性能保障北京市重点实验室, 北京 100044)

摘要: 针对滚动轴承振动信号故障特征难以自动提取和故障类别难以自动准确识别的问题, 提出一种改进集成深层自编码器 (IEDAE) 方法. 首先, 改进自编码器的损失函数并设计 3 种小波卷积自编码器; 其次, 利用区分自编码器、小波卷积自编码器等 5 种自编码器构造相应的深层自编码器, 并设计“跨层”连接以缓解深层网络的梯度消失现象, 实现对轴承振动信号的无监督预训练和有监督微调; 最后, 通过加权平均法输出识别结果, 以保证诊断结果的准确性和稳定性. 实验结果表明, 改进集成深层自编码器方法能有效地对滚动轴承进行多种工况和多种故障程度的识别, 较好地摆脱了对人工特征提取的依赖, 特征提取能力和识别能力优于现有其他方法.

关键词: 滚动轴承; 故障诊断; 深层自编码器; 集成学习

中图分类号: TH133.3

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0270

开放科学(资源服务)标识码(OSID):



引用格式: 陈志刚, 杜小磊, 王衍学, 等. 改进集成深层自编码器在轴承故障诊断中的应用[J]. 控制与决策, 2021, 36(1): 135-142.

Application of improved ensemble deep auto-encoder in bearing fault diagnosis

CHEN Zhi-gang^{1,2†}, DU Xiao-lei^{1,2}, WANG Yan-xue^{1,3}, ZHANG Nan¹

(1. School of Machine-electricity and Automobile Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China; 2. Beijing Engineering Research Center of Monitoring for Construction Safety, Beijing 100044, China; 3. Beijing Key Laboratory of Performance Guarantee on Urban Rail Transit Vehicles, Beijing University of Civil Engineering and Architecture, Beijing 100044, China)

Abstract: Considering that the difficulty of fault feature extraction of rolling bearing vibration signals and the difficulty of automatic and accurate identification of fault categories, a method based on improved ensemble deep auto-encoder (IEDAE) is proposed. Firstly, the loss function of the auto-encoder is improved and three kinds of wavelet convolution auto-encoders are designed. Then, five kinds of auto-encoders, such as discriminative auto-encoder and wavelet convolution auto-encoder, are employed to construct the corresponding deep auto-encoders, and a "cross-layer" connection is designed to alleviate the gradient disappearance of the deep network, and unsupervised pre-training and supervised fine-tuning of bearing vibration signals are realized. Finally, the recognition result is given using the weighted averaging method to ensure accurate and stable diagnosis result. The experimental results show that the IEDAE can effectively identify the bearing faults under multiple working conditions and multiple fault severities, which can get rid of the dependence on manual feature extraction and has better ability of feature extraction and recognition than other existing methods.

Keywords: rolling bearing; fault diagnosis; deep auto-encoder; ensemble learning

0 引言

滚动轴承在现代工业中起着重要作用, 自动、准确地识别滚动轴承的运行状况具有重要意义^[1]. 但是, 实际轴承振动信号易受噪声干扰, 且具有较强的非平稳性, 故障特征提取较为困难. 传统基于 ANN 和

SVM 的方法^[2-4]不能准确地诊断轴承故障^[5]. 深度学习^[6]能自动地从轴承振动信号中学习具有代表性的特征, 较好地摆脱对信号处理技术的依赖. 文献[7]利用深层降噪自编码器诊断轴承故障, 提高了隐层特征的鲁棒性; 文献[8]将轴承振动信号的时-频图像作为

收稿日期: 2019-03-10; 修回日期: 2019-10-28.

基金项目: 国家自然科学基金项目(51605022, 51875032); 北京市优秀人才培养资助项目(2013D005017000013); 北京市教育委员会科技计划一般项目(SQKM201710016014); 北京市属高校基本科研业务费专项资金项目(X18217).

责任编辑: 董久祥.

†通讯作者. E-mail: gangzi22@163.com.

卷积神经网络的输入,取得了90%以上的诊断准确率;文献[9]将轴承振动信号直接输入深层稀疏自编码器进行自动特征提取与诊断.上述研究均利用深层模型较好地实现了滚动轴承故障诊断,但同时存在以下问题:1)深层网络在有监督微调阶段,易产生梯度消失现象,从而降低诊断正确率^[10].2)由于轴承振动信号较复杂,使用单一深层模型进行故障诊断泛化能力较低^[11].

针对上述问题,本文设计了一系列具有不同特性且带“跨层”连接的深层网络,通过加权平均策略输出结果,并通过实验对所提出方法进行了验证.

1 深层自编码器

自编码器^[12](auto-encoder, AE)包括编码器和解码器,通过最小化输入和输出之间的重构误差自动完成特征提取.深层自编码器(deep auto-encoder, DAE)堆叠多个AE,采取逐层训练方法,将上一级AE的隐层输出作为下一级AE的输入,同时保证重构误差最小化,从而构成深层网络,如图1所示. DAE预训练为无监督学习过程,所需样本为无标签样本;预训练完成后,为优化网络特征,在DAE最后一层加上Softmax分类器,使用带标签样本对网络整体进行监督微调.

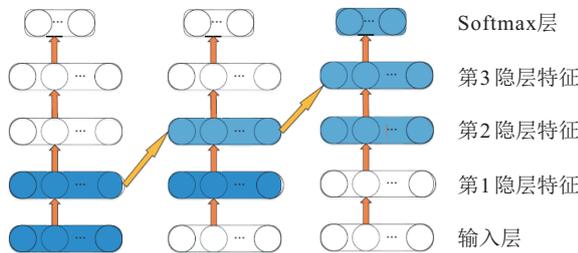


图1 3隐层DAE结构

2 “跨层”连接

“跨层”连接较好地缓解了卷积神经网络的梯度消失现象,在ResNet上具有良好的表现,因此本文将应用于DAE,使用单个“跨层”连接的深层网络结构,如图2所示.图2(a)所示的网络可分为两部分理解:前向传播和反向传播过程,定义“ \oplus ”处的传播方式如下:

1) 前向传播.设DAE有 n 个隐层,信号经第1隐层后分为两部分,一部分与DAE前向传播方式相同,为逐层向下传递;另一部分“跨层”与第 $n-1$ 个隐层的输出融合后输入到第 n 个隐层.即信号进入输入层后,在第1隐层上产生一条辅线,将其提取的特征与第 $n-1$ 个隐层提取的特征融合,输入到第 n 个隐层,即

$$H_n = (H_1, H_{n-1}), \quad (1)$$

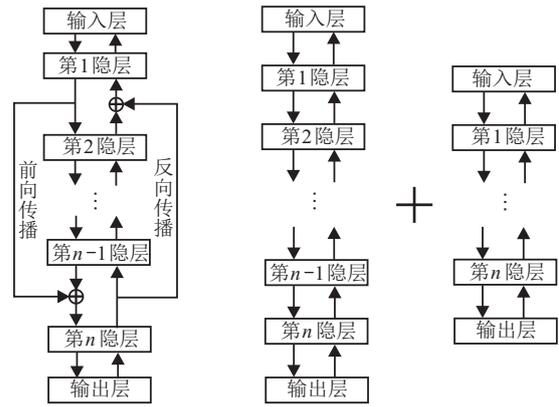


图2 DAE所拆分的两个网络

其中 H_n 、 H_1 和 H_{n-1} 分别为第 n 隐层、第1隐层和第 $n-1$ 隐层的特征向量.

2) 反向传播.反向传播时梯度向量经第 n 隐层后,分为两部分,一部分与DAE反向传播方式一致,即逐层向上传递;另一部分“跨层”与第2个隐层的输出融合并输入到第1隐层.在第 n 隐层进行反向传播时,其梯度向量分为主线 g_n^1 和辅线 g_n^2 .主线 g_n^1 的传播方式为逐层传播,经多个隐层传至 G_2 ,辅线 g_n^2 直接与 G_2 加权耦合得到 G_1 ,作为第1隐层的梯度向量,具体如下:

$$G_1 = pG_2 + qG_n. \quad (2)$$

其中: G_1 为第1隐层的梯度向量, G_2 为第2隐层的梯度向量, G_n 为第 n 隐层的梯度向量, p 和 q 为主线和辅线的梯度向量耦合比例,用 $(p:q)$ 表示.

3 改进自编码器

3.1 改进损失函数

由于最大相关熵函数对轴承背景噪声不敏感,本文以最大相关熵函数作为各深层网络的损失函数.设两个随机变量 $A = [A_1, A_2, \dots, A_m]^T$, $B = [B_1, B_2, \dots, B_m]^T$,其相关熵的近似计算如下:

$$rs_\sigma = \frac{1}{m} \sum_{i=1}^m k_\sigma(A_i, B_i), \quad (3)$$

$$k_\sigma(A_i, B_i) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{-\frac{(A_i - B_i)^2}{(2\sigma^2)}}. \quad (4)$$

其中: σ 为高斯核函数尺寸,网络的损失函数最小化可以通过最大化如下函数实现:

$$J = \frac{1}{N} \sum_{i=1}^N k_\sigma(x_i - y_i). \quad (5)$$

N 为样本数, x_i 为输入向量, y_i 为输出向量.

3.2 收缩自编码器

收缩自编码器^[13](contractive auto-encoder, CAE)利用收缩惩罚项学习信号的鲁棒性特征,其损失函数

如下:

$$J_{\text{CAE}} = -J + \lambda \|J_h(\mathbf{x})\|^2, \quad (6)$$

$$\|J_h(\mathbf{x})\|^2 = \sum_{j=1}^M \sum_{i=1}^N \left(\frac{\partial h_j}{\partial x_i} \right). \quad (7)$$

其中: $J_h(\mathbf{x})$ 是隐层单元的 Jacobian 矩阵, $\|J_h(\mathbf{x})\|^2$ 是收缩惩罚项, λ 是收缩惩罚项系数. 通过最小化 $\|J_h(\mathbf{x})\|^2$, 使 CAE 学到的隐层特征对输入的狭小变动具有鲁棒性. 收缩自编码深层网络 (contractive auto-encoder deep network, CAEDN) 的建立同 DAE, 如图 1 所示.

3.3 区分自编码器

区分自编码器^[13]利用 Fisher 惩罚项, 使网络学习到的隐层特征对信号的结构变化具有可区分性, 其损失函数如下:

$$J_{\text{Fisher}} = -J + \gamma (\text{Tr}(\mathbf{S}_w) - \text{Tr}(\mathbf{S}_b)). \quad (8)$$

其中

$$\mathbf{S}_w = \sum_{k=1}^K \sum_{n \in C_k} (\mathbf{h}_n - \mathbf{m}_k)(\mathbf{h}_n - \mathbf{m}_k)^T, \quad (9)$$

$$\mathbf{S}_b = \sum_{k=1}^K N_k (\mathbf{m}_k - \mathbf{m})(\mathbf{m}_k - \mathbf{m})^T; \quad (10)$$

$$\mathbf{m}_k = \frac{\sum_{n \in C_k} \mathbf{h}_n}{N_k}, \quad (11)$$

$$\mathbf{m} = \frac{\sum_{n=1}^N \mathbf{h}_n}{N}. \quad (12)$$

\mathbf{h}_n 为 AE 的隐层特征, N_k 为类别 C_k 的样本数, K 为故障类别数, \mathbf{m}_k 和 \mathbf{m} 分别表示类别 C_k 隐层特征的平均向量和总数据集隐层特征的平均向量, γ 为 Fisher 惩罚项系数, N 为总样本数目, $\text{Tr}()$ 为取矩阵的迹操作. \mathbf{S}_w 为内离散度, \mathbf{S}_b 为间离散度, 区分自编码器的目的是使得内离散度尽量小, 间离散度尽量大, 这对于轴承故障识别任务非常有利. 区分自编码深层网络 (discriminative auto-encoder deep network, DAEDN) 的建立同 DAE, 如图 1 所示.

3.4 小波卷积自编码器

小波自编码器^[14]使用小波激活函数, 具有更优异的特征提取性能, 但其为全连接结构, 训练时间较长; 而小波卷积自编码器^[15] (wavelet convolutional auto-encoder, WCAE) 将小波自编码器和卷积神经网络^[16] 相结合, 具有权值共享与稀疏连接优势, 所需调整的参数较少, 学习到的特征在位移和尺度上具有不变性. 设输入信号 \mathbf{x} , WCAE 隐层第 k 个节点的输出

特征为

$$\mathbf{h}_k = \psi[(\mathbf{x} * \mathbf{W}^k - \mathbf{c}^k) / \mathbf{a}^k]. \quad (13)$$

其中: ψ 为小波函数, \mathbf{W}^k 为卷积核权重矩阵, \mathbf{a}^k 和 \mathbf{c}^k 分别为隐层小波节点的尺度向量和平移向量, $*$ 为卷积符号, $/$ 为逐元素相除符号. 重构信号输出如下:

$$\mathbf{y} = \text{Sigmoid} \left[\sum_{k=1}^L \mathbf{h}_k * \mathbf{W}_T^k + \mathbf{v} \right]. \quad (14)$$

其中: L 为隐层节点个数, \mathbf{W}_T^k 为卷积核权重矩阵转置, \mathbf{v} 为偏置. 以 WCAE 为例, 其参数更新如下:

$$\begin{cases} \mathbf{W}^k \leftarrow \mathbf{W}^k - \eta \frac{\partial J_{\text{WCAE}}}{\partial \mathbf{W}^k}, \\ \mathbf{a}^k \leftarrow \mathbf{a}^k - \eta \frac{\partial J_{\text{WCAE}}}{\partial \mathbf{a}^k}, \\ \mathbf{c}^k \leftarrow \mathbf{c}^k - \eta \frac{\partial J_{\text{WCAE}}}{\partial \mathbf{c}^k}, \end{cases} \quad (15)$$

其中 η 为学习率. 小波卷积自编码深层网络 (wavelet convolutional auto-encoder deep network, WCAEDN) 的建立同 DAE, 如图 1 所示. 不同的小波函数构成不同的 WCAEDN, 将表现出不同的特征和互补的学习行为, 能有效提高模型的诊断正确率. 本文利用 3 种不同的小波函数来构造不同的 WCAEDN, 3 种小波激活函数的方程列于表 1.

表 1 3 种小波激活函数的方程

激活函数	方程
Gaussian 小波	$\psi\left(\frac{t-c}{a}\right) = \frac{t-c}{\sqrt{2\pi}a} e^{-\frac{(t-c)^2}{2a^2}}$
Morlet 小波	$\psi\left(\frac{t-c}{a}\right) = \cos\left(5\frac{t-c}{a}\right) e^{-\frac{(t-c)^2}{2a^2}}$
Mexican hat 小波	$\psi\left(\frac{t-c}{a}\right) = \left(1 - \left(\frac{t-c}{a}\right)^2\right) e^{-\frac{(t-c)^2}{2a^2}}$

3.5 加权平均法

单一深层模型对轴承进行故障识别的泛化能力较低, 而集成学习通过组合多个模型能够获得比单一模型更好的学习效果. 本文利用 5 个不同的深层网络 (CAEDN、DAEDN 和 3 个 WCAEDN) 完成轴承故障识别, 当第 i 个深层网络的分类结果为 out_i ($i = 1, 2, 3, 4, 5$) 时, 集成模型的分类输出 Y 通过将各深层网络的分类结果加权平均计算得到, 即

$$Y = \sum_{i=1}^5 w_i \text{out}_i \quad (16)$$

$$w_i = \frac{r}{\sum_{r=1}^5 r}. \quad (17)$$

其中: $r = 1, 2, 3, 4, 5$, 表示各深层网络的精度排序, r 的确定由 F_1 值的大小确定, 即

$$\begin{cases} F_1 = \frac{2PR}{P+R}, \\ P = \frac{T_P}{T_P + F_P}, \\ R = \frac{T_N}{T_P + F_N}. \end{cases} \quad (18)$$

T_P 为在正样本中被模型判断为正的个数; F_P 为在负样本中被模型判断为正的个数; F_N 为在正样本中被模型判断为负的个数; F_1 值范围在 $[0, 1]$ 之间, 反映查准率 P 和召回率 R 的信息, 0 代表最差, 1 代表最好。

综上, 本文方法流程如图3所示, 详细如下。

1) 获取不同工况的轴承振动信号样本并进行归

一化操作, 选取 70% 作为训练样本, 其余作为测试样本, 训练样本中随机选取 80% 作为无标签样本用于无监督预训练, 其余作为带标签样本用于有监督微调。

2) 采用 5 种不同的自编码器构造相应的深层网络, 并确定网络深度、各层神经元数量等参数。

3) 利用无标签样本对各深层网络进行无监督预训练, 并使用少量带标签样本对网络整体进行有监督微调, 通过 Softmax 分类器得到相应的分类结果, 由加权平均法输出最后的分类结果。

4) 使用测试样本测试模型的性能。

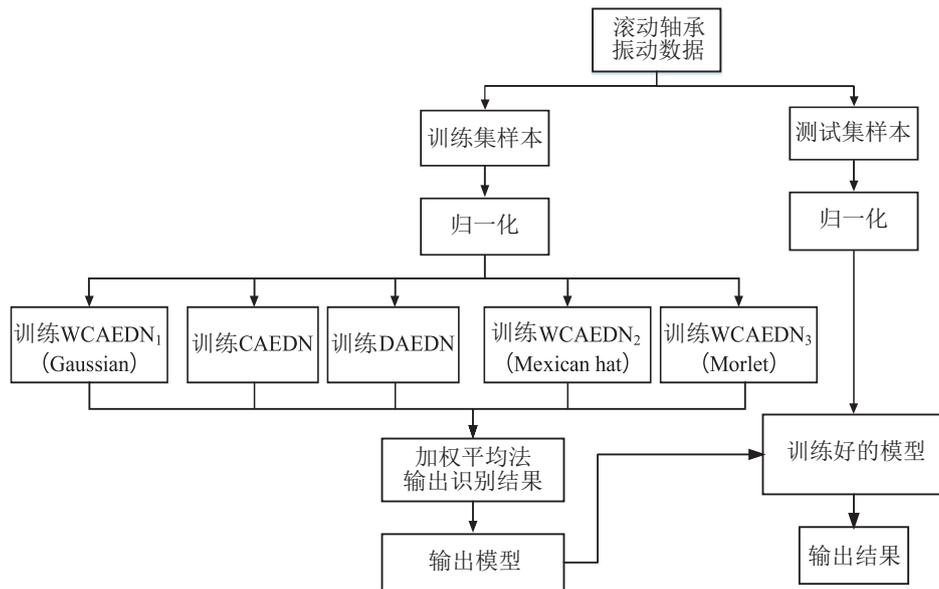


图3 故障诊断流程图

4 实验验证

4.1 实验数据初步分析

为验证本文方法的有效性, 利用轴承试验台进行分析. 试验台结构如图4所示, 由驱动器、负载、测试轴承等组成, 加速度传感器置于驱动端附近. 在试验过程中, 将轴承的外圈、内圈和滚动体分别采用电火花机床加工故障直径分别为 0.16 mm 和 0.32 mm 的环槽, 以模拟滚动轴承的不同故障程度, 采样频率设置为 12 kHz, 在 1 800 r/min、负载 1.5 hp 工况下采集轴承振动信号, 共采集不同故障类型、不同故障程度的 7 种工况, 表 2 为数据的详细介绍. 为便于网络快速收敛, 将信号归一化到 $[0, 1]$, 图 5 为归一化后的时域波形图. 可以看出, 信号受噪声干扰严重, 振动情况较为复杂, 难以区分故障类型及故障程度. 由于训练深层网络需要大量样本, 可采用文献 [17] 提出的方法对信号进行样本分割, 最后得到每种工况下 8 000 个样本, 每个样本 1 024 个采样点. 7 种工况编码样式见表 2.

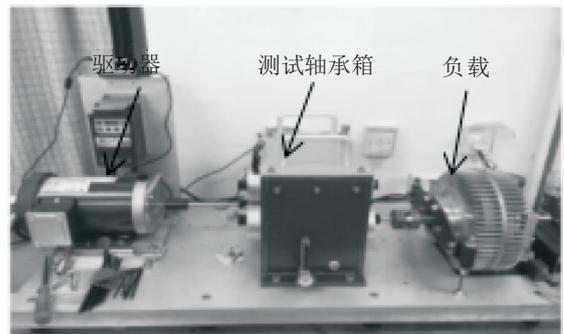


图4 轴承实验台

表2 7种轴承工况

故障位置	故障尺寸/mm	代号	故障编码
正常	0	a	1000000
滚动体+内圈	0.16	b	0100000
滚动体+外圈	0.32	c	0010000
内圈	0.16	d	0001000
内圈	0.32	e	0000100
外圈	0.16	f	0000010
外圈	0.32	g	0000001

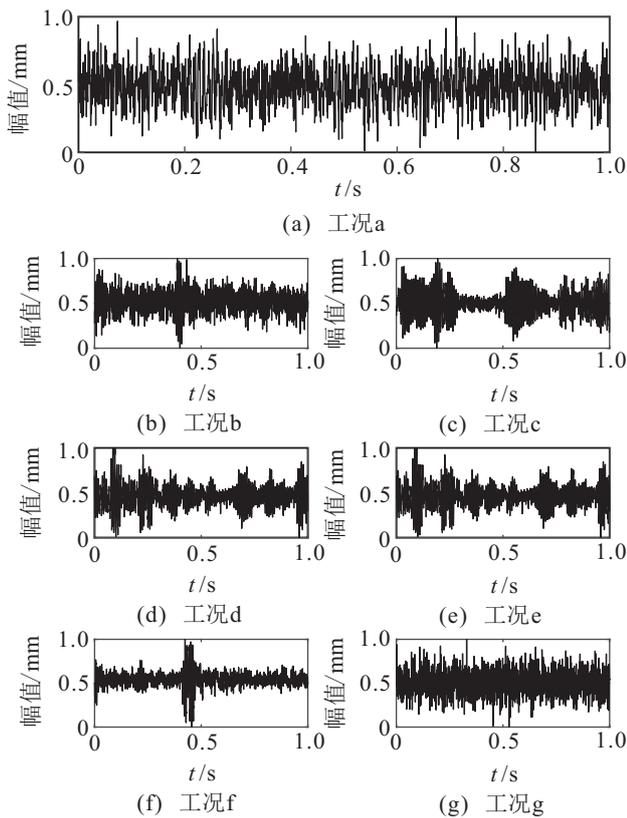


图5 轴承7种工况时域

4.2 实验结果与分析

为证明本文方法的优越性,采用 CAEDN、DAEDN 等深层模型进行分析比较,这些深层模型的输入均为归一化的 1024 维振动信号. 本文方法和其他深层模型的结构以及超参数的确定均由文献[17]中方法确定, 详细如下:

本文方法: 5 个深层网络结构均为 1024-512-256-128-64-32-7, 收缩惩罚项系数 λ 取 0.1, Fisher 惩罚项系数 γ 为 0.2, 高斯核函数 σ 为 3.5, Dropout 概率为 0.1, 每个自编码器的学习率为 0.05, 每个自编码器的迭代次数为 400, 网络整体微调迭代次数为 500 次, 主线梯度向量耦合比例 p 为 0.2, 辅线梯度向量耦合比例 q 为 0.8.

方法 2(CAEDN): 收缩惩罚项系数 λ 取 0.13, 高斯核函数 σ 为 2.5, Dropout 概率为 0.1, 每个 CAE 的学习率为 0.05, 每个 CAE 的迭代次数为 400, 网络整体微调迭代次数为 500 次, 无“跨层”连接, 结构为 1024-512-256-128-64-32-7.

方法 3(DAEDN): Fisher 惩罚项系数 γ 为 0.25, 高斯核函数 σ 为 2.8, Dropout 概率为 0.1, 每个区分自编码器的学习率为 0.05, 每个区分自编码器的迭代次数为 400, 网络整体微调迭代次数为 500 次, 无“跨层”连接, 结构为 1024-512-256-128-64-32-7.

方法 4(WCAEDN): 激活函数取 Morlet 小波, 高斯

核函数 σ 为 3.0, Dropout 概率为 0.1, 每个 WCAE 的学习率为 0.05, 每个 WCAE 的迭代次数为 400, 网络整体微调迭代次数为 500 次, 无“跨层”连接, 结构为 1024-512-256-128-64-32-7.

方法 5(3 种 WCAEDN): 将 3 种 WCAEDN 进行集成, 无“跨层”连接, 结构为 1024-512-256-128-64-32-7.

方法 6 (CAEDN + DAEDN + WCAEDN): 将 CAEDN、DAEDN 和 WCAEDN(Morlet 小波) 进行集成, 无“跨层”连接, 结构为 1024-512-256-128-64-32-7.

方法 7 (CAEDN+DAEDN + 3 种 WCAEDN): 将 CAEDN、DAEDN 和 3 种 WCAEDN 进行集成, 无“跨层”连接, 结构为 1024-512-256-128-64-32-7.

实验共进行 10 次, 图 6 列出了在每次实验中不同方法对测试样本的详细诊断结果, 表 3 列出了不同方法的平均测试精度与训练时间. 由表 3 可知, 与其他方法相比, 本文方法具有更高的识别精度 (99.08%) 和更小的标准差 (0.11), 能稳定地区分滚动轴承的不同故障类型、不同故障程度. 由图 6 可看出, 本文方法每次测试的准确率分别为 99.15%、99.08%、99.17%、99.18%、99.15%、99.16%、99.16%、99.12%、99.15% 和 99.09%, 均高于其他方法. 相对于方法 6, 本文方法由于引入“跨层”连接, 较好地缓解了网络的梯度消失现象, 加快了网络收敛, 使得诊断准确率提高了 2.46%, 训练时间也有所减少; 方法 4 中, WCAEDN 具

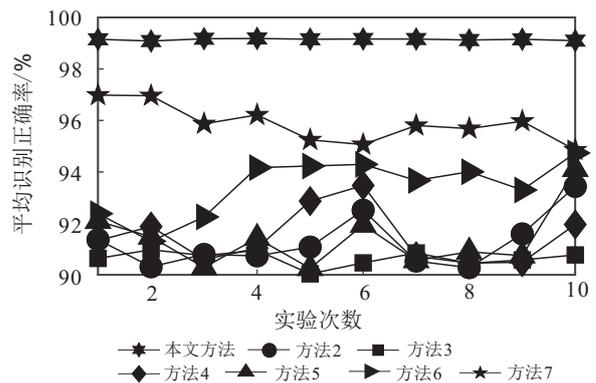


图6 不同方法的10次测试结果

表3 不同方法的诊断结果

方法	测试集平均诊断正确率±标准差/%	训练时间/s
本文方法	99.08±0.11	442.72
方法2	90.13±1.08	131.82
方法3	91.00±1.00	120.64
方法4	91.97±1.32	90.47
方法5	95.29±0.41	212.23
方法6	94.97±0.52	379.42
方法7	96.62±0.37	589.17

有时频局部化特性、稀疏连接特性和权值共享特性,网络的参数个数较少,使得网络训练时间较少。

此外,采用深度信念网络(DBN)、DAE、CNN和BP神经网络(BPNN)、支持向量机(SVM)、随机森林(RF)等方法进行分析比较,其中BPNN、SVM、RF的输入为24个时域、频域特征,详细见文献[18],主要参数如下。

DBN: 结构为1 024-512-256-128-64-32-7,学习率为0.1,动量参数0.9,RBM迭代次数为200。

DAE: 结构为1 024-512-256-128-64-32-7,学习率为0.1,动量参数0.9,AE迭代次数为200。

CNN: 结构为1个输入层、3个卷积层、3个池化层和1个输出层。输入层的尺寸为 20×1 ,第1卷积层和第2卷积层分别有6个核和12个核;2个池化层的大小均设为2,学习率为0.1。

BPNN: 结构为24-49-7,学习率为0.1,迭代次数为500。

SVM: 采用RBF核,惩罚因子和半径分别为40和0.25,由10折交叉验证方法确定。

RF: 树的总数是500,每个树的分裂节点数是4。

图7为不同方法的详细诊断结果,表4列出了测试集的平均诊断正确率。

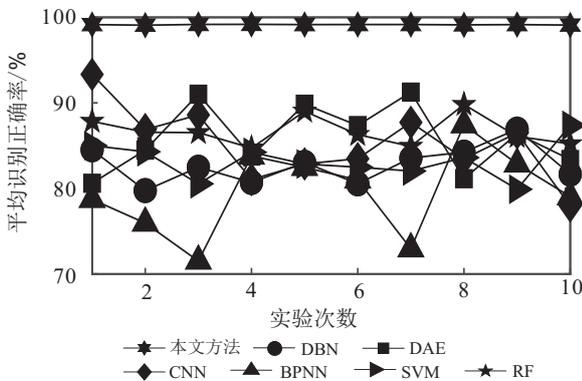


图7 不同方法的10次测试结果

表4 不同方法的诊断结果

方法	测试集平均诊断正确率±标准差/%
本文方法	99.08±0.11
DBN	82.20±2.76
DAE	86.98±2.98
CNN	84.32±3.10
BPNN	80.80±5.89
SVM	84.95±2.52
RF	86.61±2.37

由于BPNN、SVM和RF为浅层模型,性能依赖于人工特征提取,受主观因素影响较大,泛化能力低;DBN、DAE和CNN具有比浅层模型更强的非线性映射能力,但其易受噪声影响,由图7可看出,其分

类效果不如CAEDN、DAEDN和WCAEDN。图8给出了本文方法第1次测试结果的多分类混淆矩阵,可以看出,复合故障状态c的分类正确率较低。

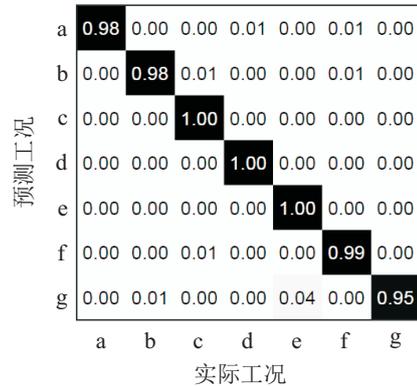


图8 第1次测试的多分类混淆矩阵

前面只考虑了基于平衡数据集的轴承故障诊断,但在实际应用中,正常样本的比例通常比故障样本的比例高;因此,本文进一步研究在不平衡数据集情况下本文方法的有效性。实验设计了4种不平衡数据集情况,比较了5种不同方法(IEDAE、无“跨层”连接IEDAE、CAEDN、DAEDN和WCAEDN)的诊断性能。每种工况下的测试样本数量为200,正常状态和各故障工况下的训练样本数量比例分别设置为2 000:2 000、2 000:1 600、2 000:1 200和2 000:1 000,实验共进行10次,5种方法的平均诊断正确率如表5所示。在第4组中,即当不平衡比例为2 000:1 000时,本文方法的平均诊断正确率为95.49%。5种方法的平均诊断正确率随着不平衡比例的增大整体上呈现下降趋势,但本文方法相较于另外几种方法在面对不同比例的不平衡样本时表现更加优异,具有更高的泛化性能。

表5 不平衡数据下不同方法的平均分类精确率 %

方法	组1	组2	组3	组4
IEDAE	99.18	97.95	96.20	95.49
无“跨层”IEDAE	96.63	96.63	93.34	91.54
CAEDN	90.13	88.04	85.95	83.85
DAEDN	91.02	88.12	85.24	82.37
WCAEDN	91.97	84.17	80.88	74.22

为定量研究基于不平衡数据集各种方法的诊断性能,本文计算了基于不平衡数据集的精确率 P 、召回率 R 和 F_1 测度,见式(18)。以组4为例,表6列出了组4中3种方法的精确率和召回率,表7显示了相应的 F_1 度量值。由表6和表7可知,组4中本文方法的 P 、 R 指标值均较高,类似的结果在其他组中也有较为明显的体现,对比结果进一步验证了本文模型在面对不平衡数据集时诊断性能的优越性和有效性,与其

他方法相比,本文方法在精确率、召回率和 F_1 测度等方面有较大的优势。

表6 组4不同方法的精确率和召回率 %

工况	IEDAE		无“跨层”IEDAE		WCAEDN	
	P	R	P	R	P	R
a	95.19	93.43	90.91	97.12	84.51	77.82
b	95.37	92.24	91.09	80.35	85.34	80.79
c	96.12	95.37	90.97	98.68	84.99	78.58
d	96.01	94.15	90.91	80.89	83.51	80.89
e	95.97	94.09	88.13	86.17	88.98	99.12
f	94.19	93.49	91.13	80.67	82.43	70.87
g	94.63	91.26	89.13	96.24	82.43	76.24

表7 组4中不同方法的 F_1 值

工况	IEDAE(F_1)	无“跨层”IEDAE(F_1)	WCAEDN(F_1)
a	0.952	0.909	0.845
b	0.954	0.910	0.853
c	0.961	0.909	0.849
d	0.960	0.90.9	0.835
e	0.959	0.881	0.889
f	0.941	0.911	0.824
g	0.946	0.891	0.824

在实际应用中,有时由于目标样本数量较少,会造成模型的性能下降,本文尝试使用迁移学习策略,构建基于辅助训练集的IEDAE模型.基本思想是将大量辅助样本用于网络预训练过程,将少量目标样本用于网络微调过程,以此实现迁移目的.为验证在目标训练样本不足情况下模型的性能,将CWRU实验室^[19]中采样频率12 kHz、转速1772 r/min的轴承振动信号作为辅助训练数据,包括表2中的7种状态,将4.1节轴承数据作为目标数据.每个信号仍然取1024个采样点,其中辅助样本集的数量为对应的目标训练样本集的5倍,目标测试样本每种状态为200个,同时,将未使用辅助训练集的IEDAE模型(即仅用少量目标训练样本训练所得的模型)作为对比.图9为两种模型诊断正确率随目标样本数目的变化结果,可以看出,随目标样本数目的增加,两种模型的诊断性能总体呈上升趋势,在目标训练样本较少的情况下,基于

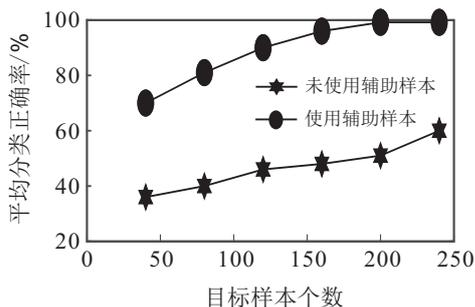


图9 两种模型的对比结果

辅助训练集的IEDAE对目标测试样本的诊断效果要好于未使用辅助训练集模型。

主、辅线比例 $p : q$ 对IEDAE的性能有一定影响,本文以 $p + q = 1$ 为约束条件,以0.1为间隔,验证不同比例对IEDAE诊断性能的影响,结果如图10所示,可以看出 $p : q = 0.2 : 0.8$ 时模型具有更优的性能。

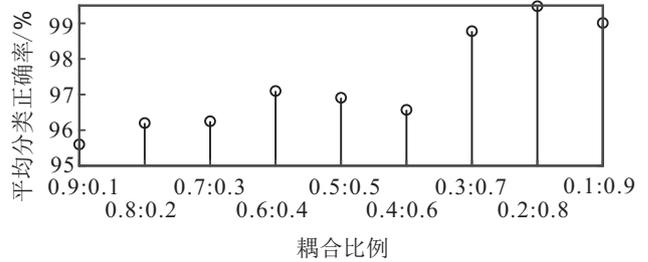


图10 主、辅线比例对IEDAE性能的影响

为直观显示IEDAE所提取的特征,以轴承外圈故障振动信号(工况f)为例,由于网络所提取的深层特征较抽象,本文只给出激活函数为Morlet小波的WCAEDN提取的第1隐层的特征图(纵坐标经归一化),如图11所示.可以看出,WCAEDN所提取的第1隐层的特征从不同角度表现原始信号,这些特征比较好地满足了Fisher判别准则^[20],更有利于最后的分类。

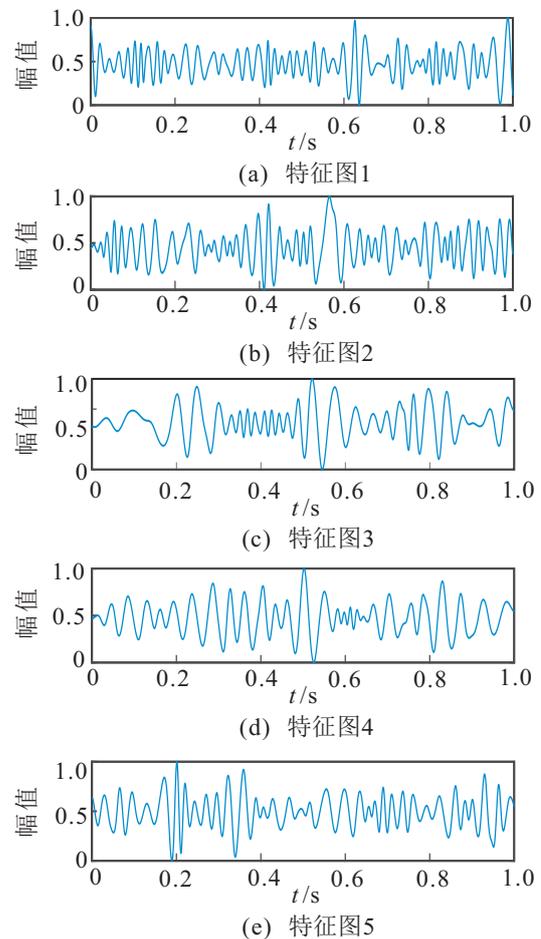


图11 WCAEDN提取的轴承外圈故障信号第1隐层特征

5 结论

本文提出了一种基于IEDAE的滚动轴承故障诊断方法,能有效地对轴承振动信号进行自动特征提取与故障识别,结论如下:

1) 将单一深层模型进行集成,充分利用了不同深层模型相互提供的互补信息,获得了比单一模型更好的学习效果。

2) 将深层小波自编码器和一维卷积神经网络结合,既使得网络具有稀疏连接和权值共享的优势,网络参数较少,学习到的特征在尺度、位移上具有不变性,又改进了自编码器的误差函数,在一定程度上增强了网络泛化性能。

3) 在深层网络的基础上,引入“跨层”连接,在一定程度上缓解了网络的梯度消失现象,并通过实验验证了“跨层”连接的有效性。

参考文献(References)

- [1] Ren H, Qu J F, Chai Y, et al. Deep learning for fault diagnosis: The state of the art and challenge[J]. Control and Decision, 2017, 32(8): 1345-1358.
- [2] Wei Z X, Wang Y X, He S L, et al. A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection[J]. Knowledge-Based Systems, 2017, 116: 1-12.
- [3] Ding X X, He Q B, Luo N W. A fusion feature and its improvement based on locality preserving projections for rolling element bearing fault classification[J]. Journal of Sound and Vibration, 2015, 335: 367-383.
- [4] Singh J, Darpe A K, Singh S P. Bearing damage assessment using Jensen-Rényi divergence based on EEMD[J]. Mechanical Systems & Signal Processing, 2017, 87: 307-339.
- [5] Shao H D, Jiang H K, Wang A, et al. An enhancement deep feature fusion method for rotating machinery fault diagnosis[J]. Knowledge-Based Systems, 2017, 119: 200-220.
- [6] Jurgen S. Deep learning in neural networks: An overview[J]. Neural Networks, 2015, 61(1): 85-117.
- [7] 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(C): 377-388.
- [8] Du X L, Chen Z G, Zhang N, et al. Faults diagnosis of rolling bearings based on synchrosqueezed S transform and deep learning[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2019(5): 90-93.
- [9] Chen Z Y, Li W H. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network[J]. IEEE Transactions on Instrumentation and Measurement, 2017, 66(7): 1693-1702.
- [10] Ma M, Xu X D, Wu J. Coupled-deep belief network based method for image recognition[J]. Journal of Xidian University, 2018, 45(5): 102-107.
- [11] Shao H D, Jiang H K, Lin Y, et al. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders[J]. Mechanical Systems and Signal Processing, 2018, 102(3): 278-297.
- [12] Du X L. Fault diagnosis method of rolling bearing based on deep learning[D]. Beijing: School of Machine-electricity and Automobile Engineering, School of Machine-electricity and Automobile Engineering, University of Civil Engineering and Architecture, 2019: 40-50.
- [13] Yuan F N, Zhang L, Shi J T, et al. Theories and applications of auto-encoder neural networks: A literature survey[J]. Chinese Journal of Computers, 2019, 42(1): 203-230.
- [14] Shao H D, Jiang H K, Zhao K W, et al. A novel tracking deep wavelet auto-encoder method for intelligent fault diagnosis of electric locomotive bearings[J]. Mechanical Systems and Signal Processing, 2018, 110(9): 193-209.
- [15] Du X L, Chen Z G, Xu X, et al. Fault diagnosis of bearing based on wavelet convolutional auto-encoder and long short term memory network[J]. Journal of Mechanical & Electrical Engineering, 2019, 36(7): 663-668.
- [16] Qu J L, Yu L, Yuan T, et al. Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network[J]. Chinese Journal of Scientific Instrument, 2018, 39(7): 134-143.
- [17] Hou W Q, Ye M, Li W H. Rolling element bearing fault classification using improved stacked de-noising auto-encoders[J]. Journal of Mechanical Engineering, 2018, 54(7): 87-96.
- [18] Qu J X, Zhang Z S, Gong T. A novel intelligent method for mechanical fault diagnosis based on dual-tree complex wavelet packet transform and multiple classifier fusion[J]. Neurocomputing, 2016, 171: 837-853.
- [19] Smith W A, Randall R B. Rolling element bearing diagnostics using the case western reserve university data: A benchmark study[J]. Signal Processing, 2015, 64(65): 100-131.
- [20] Yu H. Analog circuit fault diagnosis based on Fisher classifier[D]. Changsha: College of Electronic and Information Engineering, College of Electronic and Information Engineering, Hunan University, 2014: 54-63.

作者简介

陈志刚(1979—),男,副教授,从事智能故障诊断等研究, E-mail: gangzi22@163.com;

杜小磊(1993—),男,硕士生,从事信号处理、机器学习的研究, E-mail: 13260417530@163.com;

王衍学(1980—),男,教授,博士生导师,从事智能故障诊断等研究, E-mail: wyx1999140@126.com;

张楠(1979—),女,讲师,博士,从事动力学等研究, E-mail: zhangnan@bucea.edu.cn.