

DOI:10.3969/j.issn.1001-4551.2022.08.012

基于 Triplet Network 的小样本轴承、 齿轮故障诊断方法^{*}

谢由生, 张 军^{*}

(安徽理工大学 人工智能学院, 安徽 淮南 232001)

摘要:针对小样本条件下, 神经网络在机械设备典型部件中(轴承、齿轮)故障诊断精度不高这一问题, 提出了一种基于 Triplet Network 的机械设备典型部件(轴承、齿轮)故障诊断方法。首先, 利用短时傅里叶变换, 将原始时序信号转换为时频信号; 然后, 采用了基于 Triplet Network 搭建的模型, 从时频信号中提取了同一故障和不同故障样本的特征, 通过对比相同故障和不同故障样本特征的相似度, 对模型参数进行了优化, 达到了提取的同一故障样本特征相似度越来越高, 不同故障样本特征相似度越来越低的效果; 最后, 通过比较未知样本与已知故障样本的特征相似度, 实现了对机械设备典型部件(轴承、齿轮)的故障识别, 并采用江南大学轴承故障数据集和康涅狄格大学齿轮故障数据集, 对该故障诊断方法的有效性进行了实验验证。研究结果表明: 基于 Triplet Network 的方法在每类训练样本只有 5 个的情况下, 轴承故障识别率可以达到 68%, 齿轮故障识别率为 96.8%, 均优于传统的深度神经网络方法。

关键词: 齿轮故障诊断; 滚动轴承故障诊断; 深度神经网络; 特征相似度; 故障识别率; 时频信号; 小样本

中图分类号: TH133.33; TH132.41

文献标识码: A

文章编号: 1001-4551(2022)08-1111-07

Fault diagnosis method of small sample gear and bearing based on Triplet Network

XIE You-sheng, ZHANG Jun

(School of Artificial Intelligence, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: A fault diagnosis method of typical parts of mechanical equipment (gear, bearing) based on Triplet Network was proposed, for the problem that the fault diagnosis accuracy of deep neural network in typical parts of mechanical equipment (gear, bearing) was low under small sample conditions. Firstly, the original timing signal was converted into time-frequency signal by using short-time Fourier transform. Then, the model based on triplet network was used to extract the features of the same fault and different fault samples from the time-frequency signal. By comparing the similarity of the features of the same fault and different fault samples, the model parameters were optimized to achieve the effect that the feature similarity of the extracted same fault samples was higher and higher, and the feature similarity of different fault samples was lower and lower. Finally, fault recognition was realized by comparing the feature similarity between unknown samples and known fault samples; the effectiveness of the fault diagnosis method was verified by using the bearing fault data set of Jiangnan University and the gear fault data set of the University of Connecticut. The research results show that when there are only 5 training samples in each type, the recognition rate of bearing faults can reach 68%, and the recognition rate of gear faults can reach 96.8%, which is better than the traditional deep neural network method.

Key words: gear fault diagnosis; rolling bearing fault diagnosis; deep neural network; feature similarity; fault recognition rate; time-frequency signal; small samples

收稿日期: 2021-12-21

基金项目: 教育部产学合作协同育人项目 (201901059052, 201901228008)

作者简介: 谢由生 (1995-), 男, 江西赣州人, 硕士研究生, 主要从事机械设备故障诊断方面的研究。E-mail: xie-ys@foxmail.com

通信联系人: 张军, 男, 博士, 教授, 硕士生导师。E-mail: zhj63@163.com

0 引 言

作为现代工业中的关键设备,智能机械设备在电力、航空和工业生产等众多领域发挥着至关重要的作用。

但是智能机械设备中的典型部件,如轴承、齿轮的故障,严重威胁着智能机械设备的安全、可靠运行。故障诊断是确保这些智能机械设备长期、稳定运行的重要手段之一^[1]。

近些年,随着深度学习在自然语言处理(natural language processing, NLP)^[2]、计算机视觉(computer vision, CV)^[3]等领域取得的重大成功,众多学者把深度学习方法应用到机械设备故障诊断中,并取得了一定的成果。

CHEN L 等人^[4]提出了一种非线性频谱与堆栈降噪自编码相结合的方法,采用该方法对永磁同步电机进行故障诊断时,取得了较高的诊断准确率。KONG X 等人^[5]使用了基于深度自编码器的多集成方法,用于轴承的故障诊断,取得了较好的诊断精度,并且该方法还具有较强的泛化能力。ZARE S 等人^[6]提出了一种基于多通道卷积神经网络的故障诊断方法,采用该方法可以较好地对比风力发电机的常见故障进行诊断。赵凯辉等人^[7]采用 Inception 模型和双向长短时记忆模型(bi-directional long short-term memory, Bi-LSTM)相结合的方法,在多负载下的轴承故障诊断中取得了较好的结果,且该方法的抗干扰能力较强。

这些基于深度神经网络的故障诊断方法极大地提高了故障诊断的精度,但是这些方法的高准确率往往依赖大量的故障样本来训练模型。而在实际的生产活动中,想要获得充足的故障样本,需要花费大量的人力和物力。因此,针对小样本条件下的故障诊断是当前故障诊断领域的一大挑战。

面对小样本下故障诊断精度不高的问题,不少学者提出了解决方法。

SHAO S 等人^[8]提出了一种基于辅助分类器生成对抗网络模型(auxiliary classifier generative adversarial networks, ACGAN),用以学习真实的原始数据特征;该模型通过学习的特征,生成带有标签的伪故障样本,用来填充真实的小样本故障数据。LUO J 等人^[9]结合了深度卷积网络、条件生成对抗网络(conditional generative adversarial networks, CGAN)这二者的优点,提出了一种条件深度卷积对抗网络模型,用于生成故障样本,对小样本故障数据集进行数据增强。HAN H 等人^[10]将生成对抗网络和层叠自动编码器相结合,首先

利用生成对抗网络对小样本数据进行了增强,然后使用层叠自动编码器提取信号的特征,在对给水泵故障进行诊断时取得了较好的结果。LIU S 等人^[11]开发了一种 LOSGAN(latent optimized stable GAN)模型,在没有先验知识的情况下,自适应地增强小样本数据,并且使用分布差异距离来约束模型的优化目标,采用自注意力和谱归一化来稳定训练过程,在小样本轴承故障诊断中表现出了较优的性能。马波等人^[12]将健康状态数据反映的设备个性特征和故障机理反映的设备共性特征相结合,提出了一种基于 GAN 理论的模型,用于生成样本,对小样本故障数据进行了数据增强。

综上所述,现有提高小样本故障诊断精度的方法,多是先采用生成对抗网络生成伪样本,再对小样本故障数据进行数据增强。但是,训练一个优秀的生成对抗模型同样需要大量的样本,在实际的小样本故障诊断中存在一定的局限性。

因此,笔者提出一种基于 Triplet Network^[13]的故障诊断方法,即首先将原始故障信号转换为二维时频信号,然后对比时频信号故障样本特征的相似度,以此来改善模型提取的特征,最后通过比较未知样本与已知样本特征的相似度,实现对未知样本的故障识别。

1 模型结构与故障诊断流程

由于传统的深度学习模型很难从有限的故障样本中学习到故障的特征,基于此训练出来的模型往往是欠拟合的,故障诊断的效果较差。

笔者提出一种基于 Triplet Network 的方法,用于提高小样本故障诊断精度。

基于 Triplet Network 方法的结构图如图 1 所示。

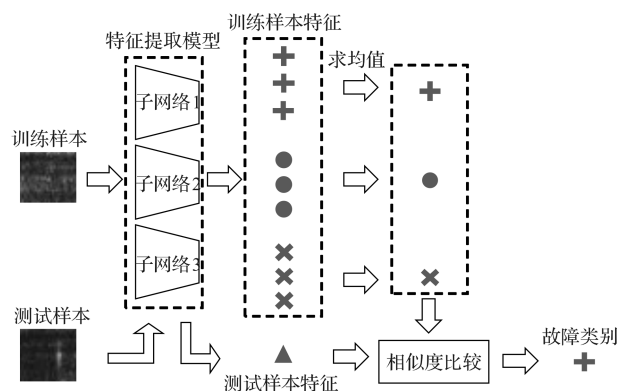


图 1 基于 Triplet Network 的方法结构图

1.1 特征提取模型

借鉴 Triplet Network 模型的思想,特征提取模型构建了一个具有 3 个共享权值的子网络模型,通过提取同一故障状态样本间和不同故障状态样本间的特征,通

过度量相同和不同故障状态样本特征的相似度,以此来优化模型抽取的特征,进而实现对不同故障的甄别。

特征提取模型结构如图 2 所示。

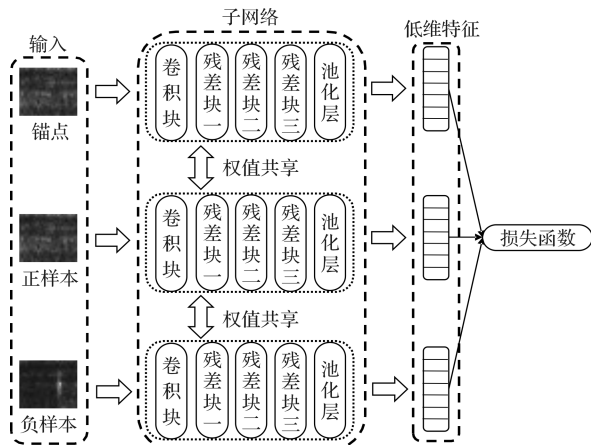


图 2 特征提取模型结构

图 2 中,模型的输入是锚点、正样本和负样本组成的样本对;锚点是从故障样本中随机选取的一个样本,正样本是与锚点同一故障状态的样本,负样本是与锚点不同故障状态的样本;输出是样本对低维特征向量间的相似度度量,即损失值;损失值通过反向传播更新模型的参数,优化模型学习的故障样本类内和类间特征,使模型提取的锚点与正样本低维特征向量的相似度越来越高,锚点与负样本低维特征向量的相似度越来越低。

模型的子网络中包含 1 个卷积块、3 个残差块和 1 个池化层,通过这些模块组成的模型,实现样本对低维特征的提取。

1.1.1 卷积块

模型的卷积块由卷积层、批归一化层、激活函数和池化层构成。

卷积层。卷积计算是现在深度神经网络的主要组成部分,卷积可以有效地提取输入对象的特征,因此也称为特征提取层。卷积中的卷积核每次提取的特征是输入对象的局部特征,但是卷积核通过滑动遍历整个输入对象,实现对整个对象的特征提取。卷积核的大小由人为设置,卷积核中的参数由梯度反向传播自动更新,解决了人工设置的繁琐。

卷积计算可以表示为:

$$X_i = W_i * X_{i-1} + b_i \quad (1)$$

式中: $*$ —卷积运算; W_i —第 i 层的卷积核; b_i —第 i 层的偏置值; X_{i-1} —第 $i-1$ 层的输出特征。

批归一化层。在模型训练过程中,参数一直处于变化状态,模型中间层的数据分布变化较大。IOFFE S 等人^[14]提出了批归一化方法,在模型中间层对数据进

行归一化处理。批归一化可以改善模型反向传播时的梯度爆炸或梯度消失问题,同时也可以起到正则化作用,一定程度上缓解了模型的过拟合现象。

激活函数。激活函数将非线性因素引入神经网络中,使神经网络可以逼近任意非线性函数。在提高模型的表示能力和学习能力上,激活函数起到了至关重要的作用。常用的激活函数有 Sigmoid 函数、Tanh 函数和 ReLU 函数等,笔者所提模型使用的激活函数有 ReLU 函数和 Sigmoid 函数。

激活函数的表达式为:

$$ReLU(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (2)$$

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

池化层。池化层对输入对象进行特征选择,降低输入对象的维度,可以大幅度地减少模型参数数据,在避免模型过拟合上有一定效果。

笔者所提模型的卷积块中采用最大池化层,残差块后的池化层为平均池化层。池化层可以表示为:

$$y = Pool(x_i) (i = 1, 2, \dots, d) \quad (4)$$

式中: y —输出; x_i —输入对象的第 i 个区域。

其中:最大池化层的 $Pool(\cdot)$ 为 $\max(\cdot)$,即取该区域中的最大值;平均池化层的 $Pool(\cdot)$ 为 $\text{mean}(\cdot)$,即取该区域的平均值。

1.1.2 残差块

2016 年,HE K^[15]提出了深度残差网络(deep residual network, ResNet),ResNet 模型在构建深度神经网络上取得了巨大成功。

笔者将 ResNet 模型中的残差结构引入到模型中,残差块结构如图 3 所示。

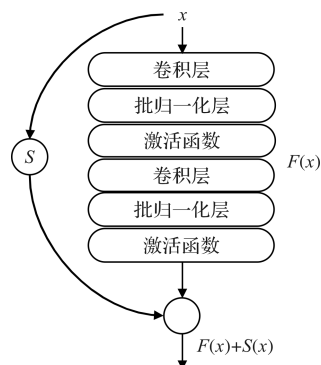


图 3 残差结构

图 3 中,输入 x 需要经过主线和支线,其中,主线由两个卷积层、批归一化层和激活函数组成;

此处 x 经过主线表示为 $F(x)$;输入 x 经过支线可表示为 $S(x)$, $S(\cdot)$ 进行何种计算需要根据该残差块

中输入 x 与 $F(x)$ 的维度决定,若 x 与 $F(x)$ 的维度相同,则 $S(\cdot)$ 不进行任何操作,直接将输入 x 输出;若 x 与 $F(x)$ 维度不相同,则 $S(\cdot)$ 为卷积核为 1 的卷积层,对输入 x 进行维度调整,使 $S(x)$ 的维度与 $F(x)$ 的维度保持一致;

输入 x 经过主线和支线后,将输出 $S(x)$ 和 $F(x)$ 相加,然后再作为后期操作的输入。

残差块可以表示为:

$$R(x) = F(x) + S(x) \quad (5)$$

1.1.3 损失函数

笔者所提模型的损失函数是度量锚点、负样本和正样本低维特征间的相似度,指导模型的优化方向,使模型提取的锚点与正样本低维特征向量的相似度越来越高,锚点与负样本的低维特征向量的相似度越来越低。

损失函数如下:

$$Loss = \max(D(a, p) - D(a, n) + margin, 0) \quad (6)$$

式中: a —锚点的低维特征; p —正样本的低维特征; n —负样本的低维特征; $D(\cdot)$ —衡量低维样本间的距离。

笔者使用的是欧式距离,距离越小相似度越高; $margin$ 是人为设定的阈值,控制模型的优化进程,若 $D(a, p) + margin > D(a, n)$,则模型继续优化,否则模型停止优化。

特征提取模型的具体结构和参数如表 1 所示。

表 1 模型详细结构参数

结构	主线	支线
卷积块	C, 11x11@32, s=2, p=1 MaxPool	—
残差块一	C, 3x3@32, s=1, p=1 C, 3x3@32, s=1, p=1	直连
残差块二	C, 3x3@64, s=2, p=1 C, 3x3@64, s=1, p=1	C, 1x1@64, s=2, p=0
残差块三	C, 3x3@128, s=2, p=1 C, 3x3@128, s=1, p=1	C, 1x1@128, s=2, p=0
池化层	AvgPool	—

C—卷积计算;11x11@32—卷积核大小为 11x11,有 32 个卷积核;s—为每次移动的步长(stride);p—填充(padding), $p=0$ 为无填充, $p=1$ 为在输入的四周围填充一个值为 0 的像素;MaxPool—最大池化;AvgPool—平均池化

模型采用的是自适应平均池化,输出为 1x1@128;支线中的直连为不进行操作,直接输出。模型中每个卷积计算后均加有批归一化和激活函数,激活函数为 ReLU 函数,平均池化层后的激活函数为 Sigmoid 函数。

1.2 故障诊断流程

模型故障诊断流程如图 4 所示。

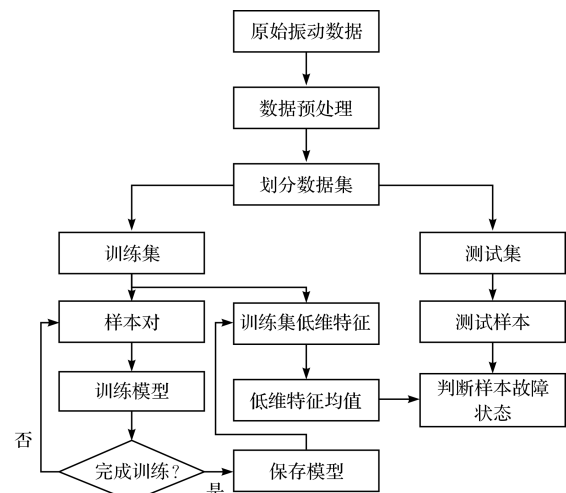


图 4 故障诊断流程

图 4 中,诊断流程分为 3 个部分:数据预处理与数据集划分、训练模型、未知样本故障诊断。

(1)数据预处理与数据集划分。笔者所提模型的输入为二维数据,训练模型前需要将原始一维数据转换为二维数据。笔者采用短时傅里叶变换,将原始时序信号转换为时频信号,并将其保存为图片,取原始时序信号中 1 024 个点转换为一张时频信号图,依次选择样本点,并将其转换为时频图,直至转换结束。

划分数据集时,按时间顺序将样本划分为训练集和测试集,如图 5 所示;

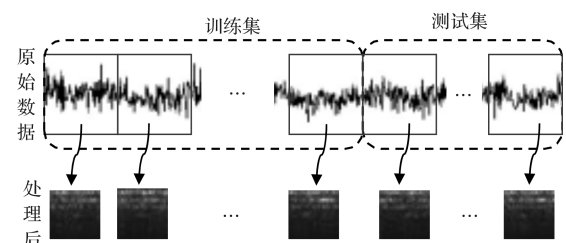


图 5 数据预处理与数据集划分

(2)训练模型。从训练集中随机选择一个样本作为锚点,随机选择一个与锚点具有相同故障状态的样本作为正样本,随机选择一个与锚点具有不同故障状态的样本作为负样本。锚点、负样本和正样本组成的样本对输入模型得到损失值,损失值经过梯度反向传播优化模型参数。如此迭代一定次数,并保存训练完成的模型;

(3)未知样本故障诊断。使用训练好的模型将训练集中各类故障样本转化为低维特征,并求各类故障样本低维特征的均值。

未知样本故障诊断如图 6 所示。

图 6 中,当诊断未知样本故障状态时,对比未知样本的低维特征到各类故障样本低维特征均值的距离,距离最小者,则判定未知样本为该故障类。使用欧式距离

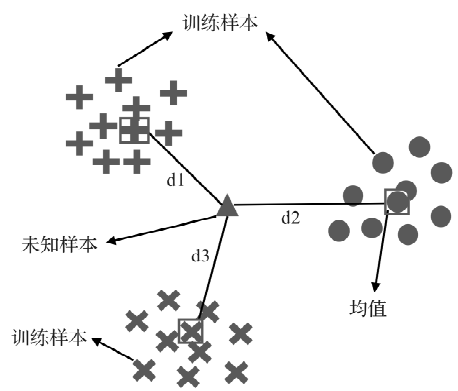


图 6 未知样本故障诊断

测量样本间低维特征的距离。

2 实验与结果分析

笔者使用江南大学轴承数据集和美国康涅狄格大学齿轮数据集,测试模型在小样本情况下的故障诊断性能。

2.1 轴承数据

江南大学轴承数据集^[16]有 4 种故障状态:健康、内圈故障、外圈故障和滚动体故障。振动信号的采集频率为 50 kHz,在转速 600 r/min、800 r/min 和 1 000 r/min 下分别采集轴承的故障信号。将转速为 800 r/min 的轴承信号进行预处理和划分数据集(预处理和划分数据集方式见第 1.2 节)。

各类故障划分数据集的结果如表 2 所示。

表 2 轴承数据集划分结果

标签	故障状态	训练集	测试集
0	内圈故障	100	30
1	健康	100	30
2	外圈故障	100	30
3	滚动体故障	100	30

2.1.1 margin 值的选择

margin 作为一个人为设定的阈值,它决定了模型是否继续优化,当故障样本的类间距离小于故障样本的类内距离加 margin 的和时,模型需要继续优化;相反,则模型停止优化。由于故障样本的类内距离小于类间距离时,模型才能做出正确的判断,所以 margin 值应该大于 0。但是,margin 值并不是越大越好,当模型可以分辨不同故障时,若再拉远不同故障类间的距离,模型的故障诊断精度不会提高。

因此,笔者在 0 到 1 之间取了多个 margin 值,测试 margin 取值对故障诊断精度的影响;从训练集中选取各类故障 30 个样本作为本实验训练集,以表 2 中的测试集作为测试样本,计算模型故障诊断精度;

训练的学习率为 0.01,每批次 64 个样本,优化器为 Adam,并且采用学习率衰减策略动态调整学习率大小,总共迭代 40 次,并且重复实验 5 次,取其平均值。

margin 值对故障诊断的影响实验结果如图 7 所示。

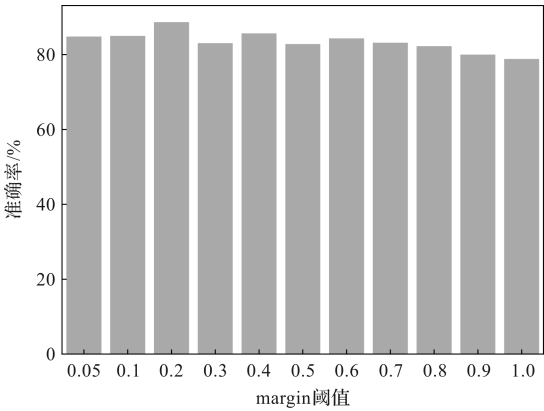


图 7 margin 值对故障诊断的影响

图 7 中,当 margin 为 0.2 时,故障诊断精度最高,所以后续实验的 margin 值均为 0.2。

2.1.2 训练样本数量对诊断精度的影响

为探究笔者所提模型在小样本下的故障诊断精度,该实验从表 2 训练集各类故障中分别随机选取了 5、10、15、30、50、80 和 100 个作为模型训练集,对比训练样本数量对模型诊断精度的影响;同时,将其与 1DCNN 模型和 2DCNN 模型作对比。

其中,1DCNN 模型是由 4 层一维卷积和全连接层组成的一维卷积神经网络,2DCNN 与笔者所提模型的结构参数相同,不同之处是 2DCNN 模型在输出低维特征后直接连接全连接层对故障进行分类。各模型使用相同的训练参数,详细参数见 2.1.1 节。

各模型共迭代 40 次,重复实验 5 次,取其平均值,得到不同模型在轴承数据中的故障诊断实验结果,如图 8 所示。

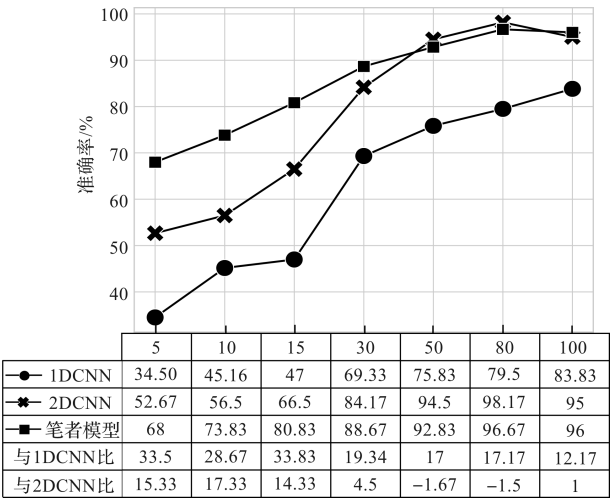


图 8 不同模型在轴承数据中的故障诊断结果

从图 8 中可知:

3 种模型整体的诊断精度随着训练样本数量的增加而提高,2DCNN 模型和笔者所提模型的诊断准确率明显高于 1DCNN 模型,笔者模型和 2DCNN 模型在每类故障样本量大于 30 时,两个模型的诊断精度较为接近,并且都大于 80% ;

在训练样本量较少时,笔者模型的诊断精度明显大于其他两个模型;在每类故障样本量为 15 时,笔者模型的故障识别率比 2DCNN 模型高 14. 33%,比 1DCNN 模型高 33. 83% ;即使是在每类故障样本只有 5 个时,笔者模型的故障识别率也有 68%,比 2DCNN 模型高 15. 33%,比 1DCNN 模型高 33. 5% 。

2.2 齿轮数据

美国康涅狄格大学齿轮数据集^[17]共有 9 种故障状态:健康、缺齿、根裂、剥落和 5 种不同程度的齿削尖。故障信号的采样频率为 20 kHz。

此处的实验选取了健康、缺齿、根裂、剥落和 1 种齿削尖共 5 类故障样本作为训练集和测试集,故障信号的预处理和数据划分方式与江南大学轴承数据集相同。

数据集划分结果如表 3 所示。

表 3 齿轮数据集划分结果

标签	故障状态	训练集	测试集	标签	故障状态	训练集	测试集
0	健康	100	30	3	剥落	100	30
1	缺齿	100	30	4	削尖	100	30
2	根裂	100	30				

从训练集的各类故障中,笔者随机选择 5、10、15、30、50、80 和 100 个样本,对比模型在不同训练集数量下的故障识别率,并将其与 1DCNN 模型和 2DCNN 模型对比。

1DCNN 模型和 2DCNN 模型结构、实验的训练参数与江南大学轴承数据集实验相同,重复实验 5 次,并取平均值,得到的实验结果如表 4 所示。

表 4 不同模型在齿轮数据中的故障诊断结果

	5	10	15	30	50	80	100
1DCNN	78.87	71.47	84	87.6	93.07	97.6	98.27
2DCNN	70.8	85.6	98.93	99.6	100	100	100
笔者模型	96.8	99.6	99.2	99.73	100	100	100
与 1DCNN 比	17.93	28.13	15.73	12.4	6.93	2.4	1.73
与 2DCNN 比	26	14	0.27	0.13	0	0	0

由表 4 可知:模型故障诊断识别率与训练样本数量呈正相关,2DCNN 模型和笔者模型明显优于 1DCNN 模型;在每类故障样本为 10 训练的 3 个模型中,笔者模型故障识别率高出 2DCNN 模型 14%,高出 1DCNN 模型 28. 13% ;甚至在每类故障样本只有 5 个的情况下,笔者模型的故障识别率依旧可以达到 96. 8% ,高出 2DCNN 模型 26% 。

综合笔者模型在江南大学轴承数据集和康涅狄格大学齿轮数据集上的表现,笔者模型在训练集样本较为充足的情况下,故障诊断精度与其他模型相差无几;但是在训练样本不足时,笔者模型的故障识别率明显优于其他模型,并且仍具有较高的故障识别率。

3 特征可视化

深度神经网络的可解释性仍然是当前的研究热点和难点,其工作原理让人难以理解。但是,其可视化部分模型的特征层可以加深对模型的理解。笔者选用在康涅狄格大学齿轮数据集中,每类故障样本为 10 训练好的 2DCNN 模型和笔者模型,把测试集样本输入模型,使用 T-SNE (T-distributed stochastic neighbor embedding)^[18] 可视化两个模型在平均池化层后输出的低维特征,其中, T-SNE 可以将高维数据映射到低维空间。

笔者将模型输出的低维特征映射到二维空间,并将其可视化,如图 9 所示。

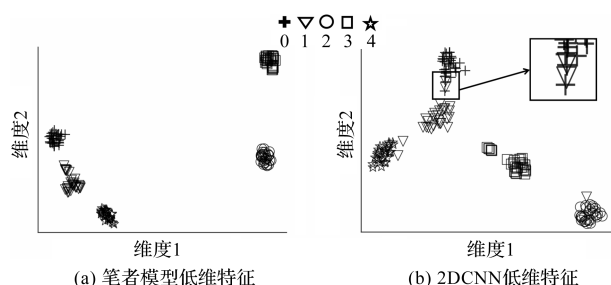


图 9 不同模型低维特征可视化对比结果

0、1、2、3、4 和 5 分别与表 3 中的标签相对应,表示不同的齿轮故障。

从图 9 中可以观察到:2DCNN 模型中的故障 1 与其他故障交织在一起,与不同故障的分界不明显,容易产生错误的故障分类结果;相对于 2DCNN 模型,笔者模型中相同故障样本的分布更加紧凑,不同故障样本的分布区别明显。

为了解 2DCNN 模型和笔者模型在齿轮数据中的分类情况,笔者绘制了混淆矩阵,如图 10 所示。

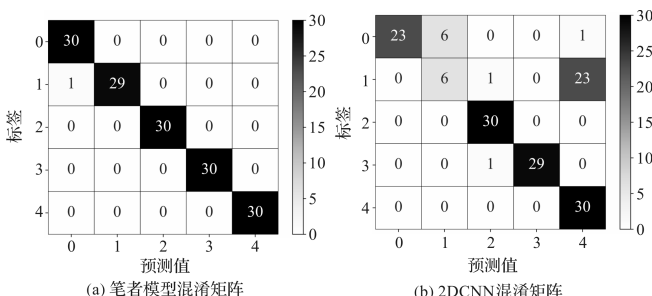


图 10 不同模型混淆矩阵对比结果

混淆矩阵的标签轴为真实的故障标签,预测值为模型的预测的故障类别。

从图 10 中可知:笔者模型在各类故障上的识别率高于 2DCNN 模型,2DCNN 模型对于故障 1 的识别误差较大。

4 结束语

笔者提出了一种在小样本数据集下机械设备典型部件的故障诊断方法,该方法可分为故障特征提取和故障诊断两个阶段。首先采用度量故障样本特征相似度的方法,来优化模型提取的特征;再通过比较未知样本与已知样本特征的相似度,实现对未知样本的故障识别。

研究结论如下:

(1) 采用短时傅里叶变换,将原始时序信号转换为时频域信号,充分挖掘原始数据信息;同时,将一维信号转换为二维图像,充分利用卷积神经网络在图像特征提取上的优势,使模型的特征提取能力得到显著提升;

(2) 利用基于 Triplet Network 搭建的模型提取故障样本时频信号低维特征,通过度量锚点与正样本低维特征和锚点,以及负样本低维特征的相似度,优化模型的特征提取策略;

(3) 通过轴承故障数据集和齿轮故障数据集验证笔者所提方法,结果表明,该方法在小样本数据下的故障诊断性能明显优于 1DCNN 和 2DCNN;在每类齿轮故障样本只有 5 个的情况下,该方法的故障识别率依旧可以达到 96.8%,比 2DCNN 高 26%,比 1DCNN 高 33.5%。

综上所述,笔者所提方法在小样本数据集下机械设备典型部件的故障诊断中具有较优的性能。

下一步,笔者考虑将该方法应用于变工况场景下旋转机械典型部件的故障诊断中,以提高该方法的泛化能力和鲁棒性。

参考文献 (References):

- [1] 雷亚国,贾峰,孔德同,等. 大数据下机械智能故障诊断的机遇与挑战[J]. 机械工程学报,2018,54(5):94-104.
- [2] YOUNG T, HAZARIKA D, PORIA S, et al. Recent trends in deep learning based natural language processing[J]. **IEEE Computational Intelligence Magazine**, 2018, 13(3):55-75.
- [3] VOULODIMOS A, DOULAMIS N, DOULAMIS A, et al. Deep learning for computer vision: a brief review[J]. **Computational Intelligence and Neuroscience**, 2018(2018):13.
- [4] CHEN L, ZHANG Z, CAO J, et al. A novel method of combining nonlinear frequency spectrum and deep learning for complex system fault diagnosis[J]. **Measurement**, 2020, 151:107190.
- [5] KONG X, MAO G, WANG Q, et al. A multi-ensemble method based on deep auto-encoders for fault diagnosis of rolling bearings[J]. **Measurement**, 2020, 151:107132.
- [6] ZARE S, AYATI M. Simultaneous fault diagnosis of wind turbine using multichannel convolutional neural networks[J]. **ISA Transactions**, 2021, 108:230-239.
- [7] 赵凯辉,吴思成,李涛,等. 基于 Inception-BLSTM 的滚动轴承故障诊断方法研究[J]. 振动与冲击, 2021, 40(17):290-297.
- [8] SHAO S, WANG P, YAN R. Generative adversarial networks for data augmentation in machine fault diagnosis[J]. **Computers in Industry**, 2019, 106:85-93.
- [9] LUO J, HUANG J, LI H. A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis[J]. **Journal of Intelligent Manufacturing**, 2021, 32(2):407-425.
- [10] HAN H, HAO L, CHENG D, et al. GAN-SAE based fault diagnosis method for electrically driven feed pumps[J]. **Plos one**, 2020, 15(10):0239070.
- [11] LIU S, CHEN J, QU C, et al. LOSGAN: latent optimized stable GAN for intelligent fault diagnosis with limited data in rotating machinery[J]. **Measurement Science and Technology**, 2021, 32(4):045101.
- [12] 马波,蔡伟东,赵大力. 基于 GAN 样本生成技术的智能诊断方法[J]. 振动与冲击, 2020, 39(18):153-160.
- [13] HOFFER E, AILON N. Deep Metric Learning Using Triplet Network[C]. International Workshop on Similarity-based Pattern Recognition. Springer, Cham, 2015.
- [14] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C]. International Conference on Machine Learning. PM-LR, 2015.
- [15] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
- [16] LI K, PING X, WANG H, et al. Sequential fuzzy diagnosis method for motor roller bearing in variable operating conditions based on vibration analysis[J]. **Sensors**, 2013, 13(6):8013-8041.
- [17] CAO P, ZHANG S, TANG J. Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning[J]. **IEEE Access**, 2018, 6:26241-26253.
- [18] VAN D M L, HINTON G. Visualizing data using t-SNE[J]. **Journal of Machine Learning Research**, 2008, 9:2579-2605.

[编辑:雷敏]

本文引用格式:

谢由生,张军. 基于 Triplet Network 的小样本轴承、齿轮故障诊断方法[J]. 机电工程, 2022, 39(8):1111-1117.

XIE You-sheng, ZHANG Jun. Fault diagnosis method of small sample gear and bearing based on Triplet Network[J]. **Journal of Mechanical & Electrical Engineering**, 2022, 39(8):1111-1117.

《机电工程》杂志: <http://www.meem.com.cn>