

DOI:10.3969/j.issn.1001-4551.2023.03.004

基于改进深度残差收缩网络的 旋转机械故障诊断*

杨正理, 吴馥云, 陈海霞

(三江学院 机械与电气工程学院, 江苏 南京 210012)

摘要: 旋转机械振动信号在多层深度学习过程中会出现退化和过拟合现象, 同时含噪数据样本也会使模型故障诊断正确率偏低, 数据样本不平衡会引起模型训练具有倾向性, 针对以上一系列问题, 提出了一种基于改进型深度残差收缩网络(DRSN)的旋转机械故障诊断方法。首先, 对多故障、长时间序列数据样本进行了矩阵化处理, 得到了模型容易接受的多维度灰度图故障样本; 针对旋转机械从正常状态到故障状态的机械老化过程, 采用了多点随机采样方法, 构建了全寿命周期数据样本, 用于后续的故障诊断; 然后, 在卷积神经网络(CNN)的基础上, 通过引入残差项、注意力机制和焦点损失函数, 构建起了多层深度残差收缩网络, 对旋转机械进行了故障诊断(其中, 残差项降低了训练过程中样本数据的特征损失, 避免了模型的退化和过拟合; 注意力机制和软阈值化自动设置噪声阈值, 降低了噪声对故障诊断精度的影响; 焦点损失函数修正了模型训练的倾向性, 提高了模型训练效率和灵敏性); 最后, 利用滚动轴承数据库样本对模型的性能进行了实例验证。研究表明: DRSN 模型在训练过程中没有出现明显的退化现象, 能够始终保持较高的训练效率和故障诊断精度, 有效避免了噪声干扰, 在不平衡数据集上修正了模型训练的倾向性。与其他模型相比较, DRSN 多层模型的平均故障诊断精度提高约 1% ~ 6%。

关键词: 滚动轴承; 卷积神经网络; 深度残差收缩网络; 软阈值化; 数据样本不平衡; 噪声干扰

中图分类号: TH133.3; TP183

文献标识码: A

文章编号: 1001-4551(2023)03-0344-09

Fault diagnosis of rotating mechanical based on improved deep residual shrinkage network

YANG Zheng-li, WU Fu-yun, CHEN Hai-xia

(School of Mechanical and Electrical Engineering, Sanjiang University, Nanjing 210012, China)

Abstract: In order to solve problems of the degradation and over fitting phenomenon of rotating machinery vibration signals in the process of multi-layer deep learning, the low accuracy of fault diagnosis caused by data sample with noise, and the model inclination in training caused by imbalance data samples, an improved deep residual shrinkage network (DRSN) based fault diagnosis method for rotating machinery was presented. Firstly, the multi-fault long time series data samples were processed by matrix, and the multi-dimensional gray scale fault samples which were easily accepted by the model were obtained. For the mechanical aging process of rotating machinery from normal state to fault state, multi-point random sampling method was used to construct the whole life cycle data sample for fault diagnosis. Then, the residual term, attention mechanism and focus loss function were introduced to construct a multi-layer deep residual contraction network to diagnose the vibration signals of rotating machinery on the basis of convolutional neural network (the residual term reduced the feature loss of sample data in the training process and avoided the model degradation and over fitting, the attention mechanism and soft thresholding automatically set the noise threshold to reduce the impact of noise on fault diagnosis accuracy, the focus loss function modified the orientation of model training and improved the efficiency and sensitivity of model training). Finally, the model was verified by using the sample of rolling bearing

收稿日期: 2022-08-07

基金项目: 江苏省高校自然科学研究面上项目(20KJB470029)

作者简介: 杨正理(1971-), 男, 江苏南京人, 硕士, 副教授, 主要从事复杂系统与计算智能方面的研究。E-mail: zhengli-yang@163.com

database. The results show that the multi-layer DRSN model has no obvious degradation phenomenon in the training process, and can maintain good training efficiency and fault diagnosis accuracy, avoid noise interference effectively, and correct the bias of model training on unbalanced data sets. Comparing with other models, the average fault diagnosis accuracy of the multi-layer DRSN model is improved by about 1% ~ 6%.

Key words: rolling bearing; convolutional neural networks (CNN); deep residual shrinkage network (DRSN); soft threshold; imbalance data samples; noise interference

0 引言

旋转机械是诸多工业领域的关键性设备,是保障工业设备持续有序运行的基础。因而,对旋转机械的运行状态进行实时检测和故障诊断已成为预测性维护维修的重要手段,是智能制造的重要组成部分。

在工作过程中,旋转机械的振动信号包含了丰富的运行状态信息和故障特征。旋转机械状态检测和故障诊断领域受到了业界人士的广泛关注。近年来,基于人工智能的传统机器学习和深度学习成为这一领域故障诊断的主流方法^[1,2]。

传统机器学习无需建立系统模型,仅利用数据驱动发掘旋转机械振动信号与设备运行状态的数据样本,构造输入特征与输出类型之间的映射关系,就能够提取其故障的通用特征,进行故障诊断。例如人工神经网络 (artificial neural network, ANN)^[3,4]、BP神经网络 (back propagation neural network, BPNN)^[5,6]、支持向量机 (support vector machine, SVM)^[7,8]、以及径向基函数神经网络 (radial basis function neural network, RBFNN)^[9,10]等。

传统机器学习算法的层数较浅,学习能力和对复杂数据特征提取的能力有限,且大数据处理能力不强。

深度学习的核心思想是采用贪婪逐层预训练、反向微调、组合应用非线性激活函数等数据驱动方法,从样本特征数据集中多角度自动提取特征向量。在状态检测、故障诊断领域,常用的深度学习网络有卷积神经网络 (CNN) 模型^[11,12]、深度置信网络 (deep belief networks, DBN) 模型^[13,14]、自动编码器 (autoencoder, AE) 模型^[15,16]等。

近年来,深度学习以其良好的学习能力和大数据处理能力在工程领域得到了广泛应用。

朱丹宸等人^[17]提出了一种改进卷积神经网络方法,将 Lenet-5 模型中单向且连续的卷积层优化成 3 个 Block 模块,并采用 L2 正则化和 Dropout 优化系统模型,防止了过拟合现象的发生,从而提取到了更精确、更完备的故障特征信息,可以在强噪声、复杂工况下,对滚动轴承进行故障诊断。PENG D 等人^[18,19]利用离

散傅立叶变换、离散余弦变换、小波变换等方法,对数据样本的特征信息进行了提取,在获得相关指标的基础上,采用 CNN 方法对旋转机械进行了智能故障诊断。温竹鹏等人^[20]针对实际工况下滚动轴承故障分布的多样性和复杂性,提出了一种改进型自编码器算法,通过求解域自适应稀疏叠加,对信号样本进行了消噪,显著提高了滚动轴承目标区域内故障诊断的精度。在实际工况下,由于缺乏部分数据标签,滚动轴承特征样本数据集的故障诊断精度不高,针对这一问题,LIANG P F 等人^[21]提出了一种区域自适应深度信念网络,采用特定域的特征信息,增强了源域和目标域信息,并将其作为模型新的输入数据,提高了滚动轴承故障诊断的精度。

旋转机械振动信号是一个高维非线性系统,构建较深层次的神经网络模型,有利于提取测量样本的特征信息,提高旋转机械的故障诊断精度。但随着深度的不断增加,网络模型容易出现退化和过拟合现象,影响了模型的学习效果。

在实际工况下,机械振动信号中存在有较多噪声,影响诊断模型的诊断精度;而且在机械振动信号样本数据集中,正常状态的数据量远多于故障状态的数据量,数据量极不平衡。不平衡的数据量容易使诊断模型训练存在倾向性,即模型在测试时容易将一些弱故障信号判定为正常,从而降低了诊断模型的灵敏度^[22]。

为了解决这些问题,笔者结合机械振动信号的数据特点,提出一种基于改进型深度残差收缩网络 (DRSN) 的旋转机械振动信号故障诊断方法。

首先,笔者对旋转机械从正常运行状态到故障状态的过程数据进行预处理,构建了多维度矩阵化灰度图故障样本,以及全寿命周期故障样本;然后,将残差项、注意力机制和焦点损失 (focal loss, FL) 函数引入到 CNN,构建多层 DRSN 模型 (残差项可避免模型在训练过程中可能出现的退化和过拟合现象;注意力机制和软阈值化可降低样本数据噪声对故障诊断精度的影响;FL 函数可改善机械振动信号的不平衡问题,修正模型训练的倾向性),形成改进型 DRSN;最后,通过实例验证上述模型在各方面性能的先进性。

1 深度残差收缩网络(DRSN)模型

1.1 深度学习基本模型

深度学习基本模型能够有效地提取大数据样本的复杂特征,并且具有良好的训练效果。

笔者以 CNN 为深度学习基本模型,以构建 DRSN。CNN 模型主要由输入层、卷积层、ReLU 层、池化(Pooling)层、全连接层和输出层构成。

输入层主要是对输入样本进行矩阵化处理,形成使模型容易接受的图片类型数据,并进行降维、归一化等预处理。在实际深度学习模型中,卷积层与 ReLU 层共同作用,通过多个卷积核对输入数据进行特征提取和保存,得到多通道特征图,从而将低维度数据集映射到高维度数据空间;

Pooling 层对卷积层输出的特征图进行特征选择和信息过滤,常用的池化函数有 Max Pooling 和 Average Pooling,其主要作用是压缩数据、减少模型过拟合现象;

全连接层是由许多神经元组成的平铺结构,等价于传统前馈神经网络中的隐含层,位于 CNN 的末端,层中的每一个节点都与前层的节点互连,对前层网络所提取的数据特征进行整合,并将输出值传送给输出层的分类器,例如 SoftMax 函数,输出分类结果。

CNN 基本模型如图 1 所示。

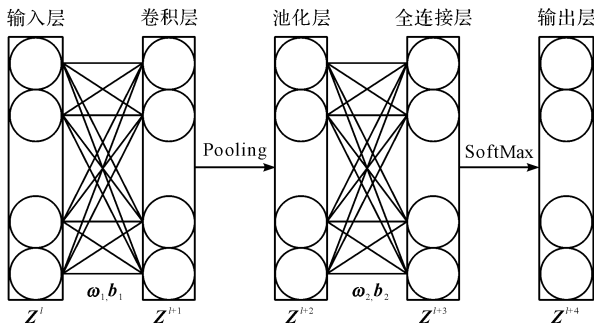


图 1 CNN 基本模型

$Z^l, Z^{l+1}, Z^{l+2}, Z^{l+3}, Z^{l+4}$ —输入层、卷积层、池化层、全连接层和输出层; ω, b —权重向量、偏置向量

在 CNN 模型中,数据在各层的前向传播函数表达式为:

$$Z^{l+1} = \text{ReLU}(\omega_1 Z^l + b_1) \quad (1)$$

$$Z^{l+2} = \text{Pooling}(Z^{l+1}) \quad (2)$$

$$Z^{l+3} = \omega_2 Z^{l+2} + b_2 \quad (3)$$

$$Z^{l+4} = S(Z^{l+3}) \quad S(x) = \frac{1}{(1 + e^{-x})} \quad (4)$$

假设输入 CNN 模型的数据样本标签为 y ,则损失函数的计算公式为:

$$J(\omega, b, Z, y) = \frac{1}{2} \| Z^{l+4} - y \|^2 \quad (5)$$

在对模型梯度进行计算时,CNN 模型各层间数据的依赖公式为^[23]:

$$\delta^{l+4} = \frac{\partial J(\omega, b, Z, y)}{\partial Z^{l+3}} = (Z^{l+4} - y) \odot S''(Z^{l+3}) \quad (6)$$

$$\delta^{l+3} = (\omega_2)^T \delta^{l+4} \odot S'(Z^{l+3}) \quad (7)$$

$$\delta^{l+2} = \text{Unsample}(\delta^{l+3}) \odot S'(Z^{l+3}) \quad (8)$$

$$\delta^{l+1} = \delta^{l+2} \text{rot}180(\omega_1) \odot S'(Z^{l+3}) \quad (9)$$

式(6~9)中:Unsample()函数的功能是将 δ^{l+3} 中各子矩阵的所有池化区域值求平均值后,再放回还原后的子矩阵位置; \odot 的功能是将两个矩阵对应位置元素进行相乘运算; $S()$ 为 sigmoid 激活函数,即 $S(x) = 1/(1 + e^{-x})$;rot180()函数的功能是将卷积核旋转 180°。

综合以上分析可知,数据在 CNN 模型各层间后向传播时,模型参数的迭代公式为:

$$\omega_2 = \omega_2 - \lambda \delta^{l+3} (Z^{l+2})^T \quad (10)$$

$$b_2 = b_2 - \lambda \delta^{l+3} \quad (11)$$

$$\omega_1 = \omega_1 - \lambda \delta^{l+1} \cdot Z^l \quad (12)$$

$$b_1 = b_1 - \lambda \sum_{u,v} (\delta^{l+1})_{u,v} \quad (13)$$

式中: λ —学习率; $\sum_{u,v} (\delta^{l+1})_{u,v}$ —求取误差向量,即对 δ^{l+1} 的各个子矩阵的项分别进行求和。

在 CNN 模型的训练过程中,为了得到良好的结果,笔者需要不断地对前向传播和后向传播的参数进行微调。

1.2 DRSN

当深度学习基本模型的网络层数越多时,其非线性表达能力就越强,模型的学习能力也就越强,能够提取到的样本数据特征就越复杂。但是,随着模型网络层数的增多,非线性表达函数很难表示数据间的恒等映射关系,使得模型在训练过程中容易出现退化和过拟合现象^[24]。另外,在实际工况下,旋转机械的振动信号中包含大量的噪声和无关特征干扰,影响了模型的故障诊断精度。

因此,为了解决上述问题,笔者提出了深度残差收缩网络(DRSN)模型。具体阐述如下:(1)当深度残差网络反向传播,进行模型训练时,不但其损失能够通过各层实现反向传播,而且恒等映射也能通过残差项,便利地实现反向传播;(2)通过注意力机制和软阈值函数,可以在 DRSN 模型内实现自动软阈值化,自适应地消除数据样本中的噪声和无关特征干扰,提高模型的故障诊断效果^[25,26]。

假设需要求解的模型的映射为 $H(Z^l)$,现将这一

求解目标转化成求解模型的残差映射函数 $F(\mathbf{Z}^l)$, 即 $F(\mathbf{Z}^l) = H(\mathbf{Z}^l) - \mathbf{Z}^l$ 。在 DRSN 模型中, 软阈值化比 ReLU 更加灵活地设置样本特征的取值区间。更为重要的是, 样本阈值在引入注意力机制下, 可以根据每个样本的具体特征实现单独设置, 即输入至网络模型的每一个样本都可以有一组自己的独特阈值, 所以 DRSN 可以适应样本中包含多种复杂噪声的情况。

DRSN 模型的局部如图 2 所示。

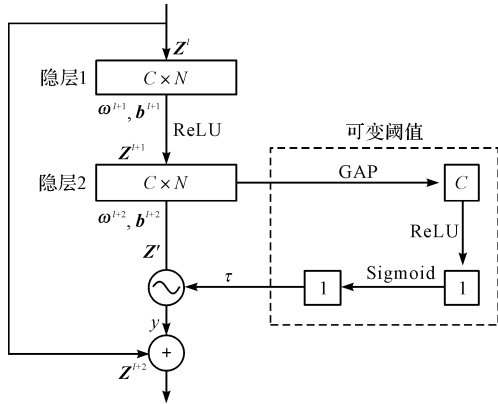


图 2 DRSN 模型局部

图 2 中, 输入样本矩阵 \mathbf{Z}^l 大小为 $C \times N$, \mathbf{Z}^l 经过隐层 1 和 ReLU 函数处理后得到 \mathbf{Z}^{l+1} , 并输入到隐层 2。在隐层 2 中建立子网络, 通过学习得到可变阈值 τ , 并经软阈值化后, 再将其与残差项 $F(\mathbf{Z}^l)$ 进行求和, 最后得到输出 \mathbf{Z}^{l+2} 。

图 2 中各层输出的结果分别为:

$$\mathbf{Z}^{l+1} = \text{ReLU}(\omega^{l+1} \mathbf{Z}^l + b^{l+1}) \quad (14)$$

$$\mathbf{Z}' = \omega^{l+2} \text{ReLU}(\omega^{l+1} \mathbf{Z}^l + b^{l+1}) + b^{l+2} = \omega^{l+2} \mathbf{Z}^{l+1} + b^{l+2} \quad (15)$$

$$y = \begin{cases} \mathbf{Z}' - \tau & \mathbf{Z}' > \tau \\ 0 & -\tau \leq \mathbf{Z}' \leq \tau \\ \mathbf{Z}' + \tau & \mathbf{Z}' < -\tau \end{cases} \quad (16)$$

$$\mathbf{Z}^{l+2} = y + F(\mathbf{Z}^l) \quad (17)$$

式(16)对应每一个维度下, \mathbf{Z}' 与该维度下的阈值 τ 进行比较后得到的软阈值化结果。

1.3 数据不平衡处理

为解决旋转机械振动信号数据集的不平衡和提高模型故障诊断的灵敏性。笔者将 FL 函数引入到 DRSN 模型中, FL 函数的表达式^[27,28]为:

$$\begin{cases} \text{FL}(p_t) = -\alpha_t(1-p_t)^\gamma \ln(p_t) \\ p_t = \begin{cases} p & x=1 \\ p-1 & x=0 \end{cases} \\ \alpha_t = \begin{cases} \alpha & x=1 \\ \alpha-1 & x=0 \end{cases} \end{cases} \quad (18)$$

式中: p —样本标签概率; t —样本类别; γ —样本实际标签。

神经网络的训练过程主要包括 FL 函数计算和梯度更新。旋转机械振动信号数据集的严重不平衡必然会造成模型更加关注正常样本的训练情况, 训练结果自然具有倾向正常状态的判定趋势。

$\alpha(0 \leq \alpha \leq 1)$ 的作用是在模型训练过程中, 调节样本的权重参数, 即调节样本中正常状态数据量和故障状态数据量的占比, 以改善真实机械样本中, 正常样本在数量上占绝大多数而导致的平衡, 从而改善模型训练倾向正常判定的问题。

另外, 为了提高模型在故障诊断时的灵敏性, 笔者引入 $(1-p_t)^\gamma$ 作为调节因子, 以解决大量正常样本可能导致 FL 值缓慢下降, 结果无法收敛到最优的问题, 从而提高模型在故障诊断时的综合灵敏性。

2 旋转机械故障诊断模型

在卷积神经网络(CNN)基础上, 笔者通过引入残差项、FL 函数、注意力机制, 构建起多层深度残差收缩网络(DRSN)模型, 用以对旋转机械进行故障诊断。

旋转机械振动信号故障诊断模型结构如图 3 所示。

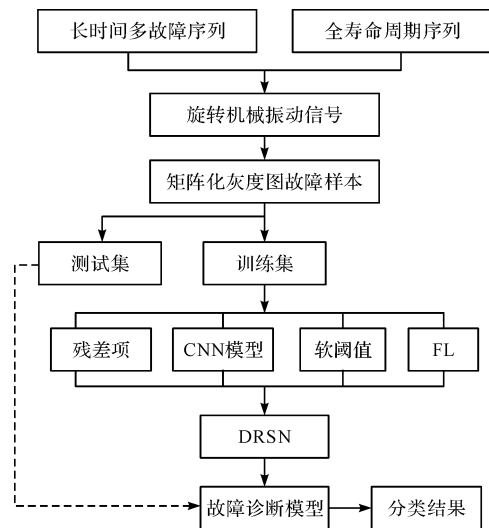


图 3 旋转机械振动信号故障诊断模型

图 3 中的恒等映射是 DRSN 的核心贡献。

残差项降低了模型训练的难度, 避免了模型退化和过拟合现象的发生;

FL 函数解决了模型训练的倾向性, 提高了模型故障诊断的灵敏性;

注意力机制下的软阈值化可使模型, 在面对含有复杂、多样噪声的机械振动信号时, 仍然具有良好的故障诊断精度。

图 3 中,笔者首先采用样本数据集,构建旋转机械振动信号的多故障、长时间序列和全寿命周期信号样本;然后将样本进行矩阵化,并构成灰度图样本(灰度图样本分为训练集和测试集),训练集输入到由残差项、FL、软阈值和多层 CNN 构建的改进型 DRSN 中,训练和优化旋转机械振动信号故障诊断模型;最后,将测试集代入旋转机械故障诊断模型中,根据故障分类的灵敏性、正确率等性能,衡量模型的有效性。

3 故障诊断实验

笔者选用美国凯斯西储大学轴承数据集和西南交通大学的轴承数据集作为原始数据样本,分别构成多故障长时间序列数据样本和全寿命周期序列数据样本^[29,30],对上述模型的各方面性能进行验证。

3.1 多故障、长时间序列数据样本

笔者选择美国凯斯西储大学轴承数据库中,采样频率为 12 kHz 的驱动端轴承故障数据。其中,故障直径为 0.178 mm,故障负载为 0 HP。

笔者构建该数据集的 5 种故障及其对应的数据长度,如表 1 所示。

表 1 机械振动故障样本数据集

故障代号	故障位置	时间序列长度
F1	内圈故障	2×10^4
F2	外圈故障	2×10^4
F3	保持架故障	2×10^4
F4	滚动体故障	2×10^4
F5	正常运行状态	2×10^4

3.1.1 故障样本预处理

笔者采用 Z-score 模型,对旋转机械振动信号多故障、长时间序列数据样本进行标准化处理,其处理结果如图 4 所示。

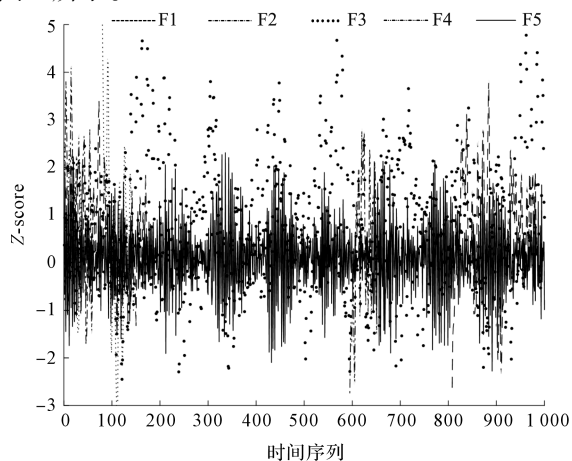


图 4 数据样本标准化结果

Z-score 模型的标准化方法是对样本数据按比例进行缩放,并使缩放结果能够落在特定区间内。

假设样本为 $x(x_1, x_2, \dots, x_k, \dots)$, 其均值和方差分别为 μ, σ , 则 Z-score 模型的标准化公式为:

$$z_k = \frac{x_k - \mu}{\sigma} \quad (19)$$

将标准化后长时间序列数据样本矩阵化成 20×20 的灰度图数据形式(灰度图的数据形式有利于输入深度学习模型)。

笔者选择 F1 样本进行预处理,其灰度图故障样本如图 5 所示。

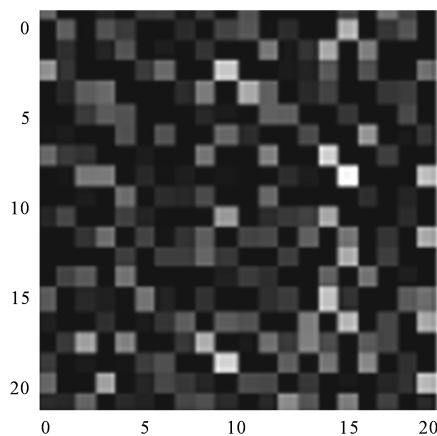


图 5 样本 F1 的灰度图故障样本

在算例的仿真过程中,为了模拟实际工况下的含噪机械振动信号,笔者在原始数据样本上添加不同方差的高斯白噪声。

此处,笔者定义两种等级噪声:(1)噪声等级 1,噪声方差 $\sigma = 0.01$, (2)噪声等级 2,噪声方差 $\sigma = 0.02$ 。

笔者分别从 F1 ~ F5 样本中,截取 1 200 个数据组成 6 000 个样本集,重复以上过程,构建机械振动信号多故障长时间序列数据样本集合。

3.1.2 模型构建与训练

笔者构建的 DRSN 模型如图 6 所示。

图 6 中的 DRSN 模型包含 10 个隐层,每 2 层引入 1 个残差模块和 1 个软阈值化。

机械振动信号经数据预处理后作为模型输入,其样本尺寸为 20×20 ,输入层参数为 20×20 -Conv-10,基础模型为 CNN,通道为 10。

在模型输入层,输入样本经 Pooling 操作和 ReLU 函数处理后,提取出尺寸为 10×10 的样本特征,作为隐层的输入数据。前 4 个隐层的参数相同,均为 10×10 -Conv-10。第 4 个隐层输出数据经 Pooling 操作,将样本特征尺寸压缩至 5×5 ,同时将通道数增加到 20 个;后 6 个隐层参数也是相同的,均为 5×5 -Conv-20。

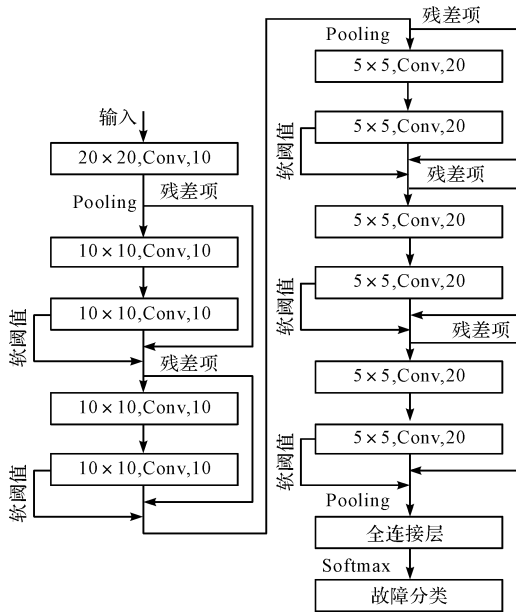


图 6 DRSN 模型结构

第 10 个隐层输出的样本特征经 Pooling 操作和 ReLU 函数处理后,将 20 个通道的样本特征展开,输入到全连接层,再经分类函数 SoftMax 处理,最后得到分类结果。

在训练 DRSN 模型时,每组输入 32 个数据样本。当迭代次数 (Epochs) 达到 60 时,DRSN 模型的训练均方误差 (mean-square error, MSE) 从 0.54 下降到 0.023,机械故障诊断正确率从 75.4% 上升到 99.6%。

2 个评价参数的变化曲线如图 7 所示。

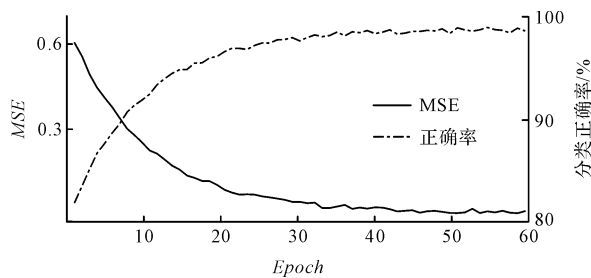


图 7 模型训练误差和正确率变化曲线

DRSN 模型训练完成后,笔者将噪声等级 1 和噪声等级 2 的白噪声分别输入到经过数据预处理(3.1.1 节)后的多故障长时间序列集中,构成两组含噪数据样本集(各 3 200 个,分别用 N1 和 N2 表示),以此作为模型的测试样本。

含噪样本故障诊断测试结果如图 8 所示。

从图 8 中数据可以看出:(1)含噪样本 N1 和 N2 的故障诊断正确率分别为 99.6% 和 99.2%,这说明,当样本噪声越大时,对模型的故障诊断正确率影响越大;(2)上述模型在较强的噪声干扰下,仍然具有较高的故障诊断正确率。

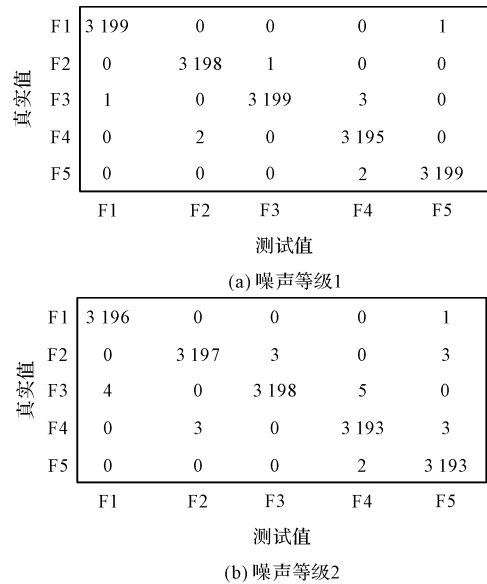


图 8 含噪样本故障诊断结果

笔者将多故障、长时间序列数据样本分别输入到 DRSN 模型、深度残差网络 (DRN) 模型和不含残差收缩功能的 CNN 模型,并分别进行 60 次迭代训练,得到的训练 MSE 分别为 0.018 6%、0.027 8%、0.031 7%。

笔者绘制了 DRSN 模型、深度残差网络 (DRN) 模型和不含残差收缩功能 CNN 模型的训练误差曲线,如图 9 所示。

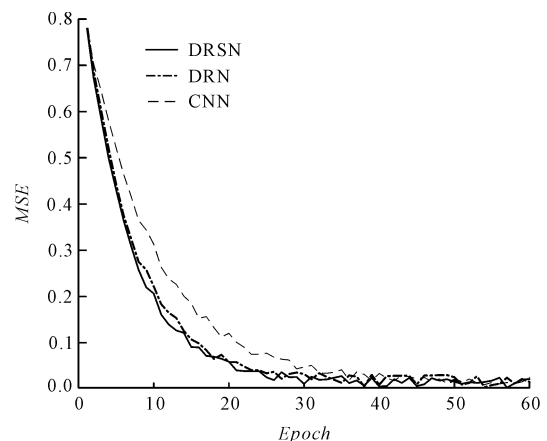


图 9 不同模型训练误差曲线

图 9 中曲线说明:(1)与 DRN 模型和 CNN 模型相比,DRSN 模型的训练误差下降得更快,并能较快地收敛到较低的训练误差;(2)虽然在 CNN 模型中引入了多个模块,但并没有过多影响 DRSN 模型快速性的。

3.2 全寿命周期数据样本

笔者选择西南交通大学轴承数据库中的外圈故障和保持架故障时数据集,作为全寿命周期序列数据样本^[31],运行条件为 37.5 Hz/11 kN。

该数据集采样了旋转机械在正常状态下,分别运行 5 h 39 min 和 8 h 53 min 后发生外圈故障和保持架故障

的全过程状态数据,截取了故障发生前 200 个采样点以内的振动信号样本数据,将每个样本数据所对应的长时间序列划分(按定长)为多个小样本集,并从中随机无放回地抽取小样本,构造了包括 200 维水平振动信号,以及 200 维垂直振动信号的全寿命周期故障样本。

全寿命周期故障样本经过数据预处理(3.1.1 节)后,笔者将模型的输出节点修改为 2,对上述样本进行故障诊断。

经过 60 次 Epochs 后,训练 MSE 降低到 0.049 7,故障诊断正确率上升为 99.5%。这说明,数据集经过特定方法处理后,DRSN 模型能够准确地判断出数据样本的特征指标所对应的故障类型。

3.3 对比实验

为了全面检验 DRSN 模型的性能,在相同条件下,笔者对比 DRSN 模型、传统 DRN 模型、CNN 模型、DBN 模型、ANN 模型和 SVM 模型的故障诊断性能。

笔者对 400 维样本输入、输出节点为 6 的数据集 1 和输出节点为 2 的数据集 2 的故障诊断,并将两种噪声等级的白噪声分别引入到数据集 1 和数据集 2,然后针对不含噪声样本和含两种不同噪声等级的样本,分别采用上述 6 种不同模型进行故障分类。

故障诊断正确率如表 2 所示。

表 2 故障诊断正确率对比(%)

模型	数据集 1			数据集 2			平均值
	原始	N1	N2	原始	N1	N2	
DRSN	99.3	98.9	98.0	99.7	99.1	98.9	99.0
DBN	96.6	95.7	95.6	97.9	97.8	96.9	96.7
DRN	98.4	97.7	97.0	98.9	98.6	98.1	98.1
CNN	97.6	97.1	96.9	98.1	97.8	97.5	97.5
ANN	91.3	91.2	90.7	94.6	93.8	92.6	92.4
SVM	90.7	89.6	88.3	93.6	94.0	93.2	91.6

从表 2 中数据可以看出:在不同数据集和噪声等级下,机器学习算法的故障诊断正确率最低,而 DRSN 模型故障诊断正确率最高。

这说明,深度学习算法比机器学习算法具有更好的提取样本特征和学习能力;而基于深度学习的模型,在引入残差项、注意力机制和损失函数等模块后,其所构建的 DRSN 模型,在样本特征提取、样本去噪、避免模型过拟合、提高故障诊断等性能上均有所提高。

与其他模型相比,DRSN 模型的故障平均正确率提高了约 1%~6%,同时其整体性能也得到了一定程度的改善。

3.4 不平衡数据样本

在实际工况下,旋转机械振动信号中的正常状态

数据量远大于故障状态数据量,因此,机械振动信号数据集是极不平衡的。

在面对不平衡数据时,为了检验改进型 DRSN 模型的适应能力,笔者在原始数据集中分别抽取正常状态数据和故障状态数据,分别构成 2 400 : 2 400、2 800 : 2 000、3 200 : 1 600、3 600 : 1 200 等 4 种比重的数据集,然后分别利用模型(3.3 节中),对 400 维的样本输入、输出节点为 6 的数据集进行故障诊断。

在数据不平衡条件下,模型的故障诊断正确率如图 10 所示。

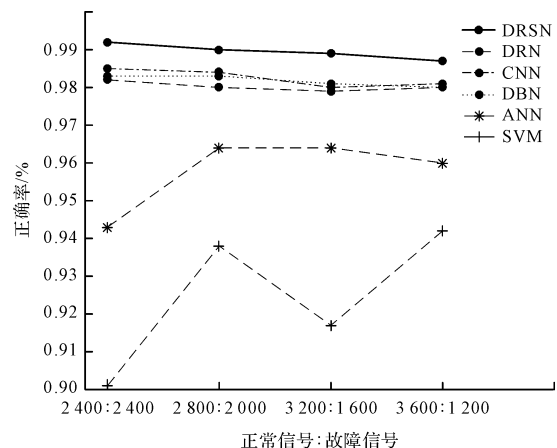


图 10 数据不平衡条件下故障诊断正确率对比

从图 10 中数据可以看出:(1)随着正常状态样本的增多,机器学习算法受到数据不平衡的影响要大于神经网络模型;(2)DRSN 模型以神经网络为基础,且有针对性地对数据不平衡进行了改进,所以其故障诊断正确率能始终保持在较高水平,且随着数据不平衡度的变化,其波动相对较小。

笔者采用数据状态比重为 3 600 : 1 200 的数据集对 DRSN 模型、不引入 FL 函数的所建模型,分别进行训练,再分别采用数据状态比重为 2 400 : 2 400、2 800 : 2 000、3 200 : 1 600 的数据集进行了测试,最后得到了两种模型的故障诊断混淆矩阵,如图 11 所示。

实际种类	2 400 : 2 400	2 386	2 414
	2 800 : 2 000	2 790	2 010
	3 200 : 1 600	3 195	1 605
		正常	故障
		诊断种类	

(a) 不引入 FL 函数的本文模型

实际种类	2 400 : 2 400	2 396	2 404
	2 800 : 2 000	2 798	2 002
	3 200 : 1 600	3 199	1 601
		正常	故障
		诊断种类	

(b) 本文模型

图 11 两种模型的混淆矩阵

图 11 中数据表明:(1)DRSN 模型通过引入 FL 函数,修正了模型故障诊断的倾向性,提高了模型故障诊断的精确度和灵敏性;(2)DRSN 模型能够更好地适应实际正常状态下机械振动信号占绝大多数情况,因而具有较好的实用性。

4 结束语

在多层深度学习过程中,随着深度的不断增加,网络模型容易出现退化和过拟合现象,影响了模型的学习效果。为此,笔者提出了一种基于改进深度残差收缩网络(DRSN)模型的旋转机械故障诊断方法。

首先,笔者针对旋转机械从正常运行状态到故障状态的过程数据进行预处理,构建了多维度矩阵化灰度图故障样本,以及全寿命周期故障样本;然后,将残差项、注意力机制和焦点损失(FL)函数引入到 CNN,构建了多层 DRSN 模型;最后,通过实例验证上述模型在各方面性能的先进性。

研究表明:

(1)针对多故障、长时间旋转机械一维振动信号,将数据矩阵化为灰度图故障样本,便于模型训练;

(2)针对从正常状态到故障状态的机械老化过程,对多个采样点进行随机采样,构成了能够表现机械老化过程特征的全寿命周期故障样本,可将其应用于旋转机械的故障诊断;

(3)针对旋转机械振动信号这类典型的不平衡数据集分类问题,在 DRSN 模型中引入 FL 函数,改善了模型训练具有倾向正常分类状态的趋势;与其他模型相比,DRSN 模型的故障诊断平均正确率提高了约 1%~6%,同时其训练效率和模型故障诊断灵敏度也得到了提高;

(4)在 DRSN 模型中引入残差项,降低了多层模型训练过程中的样本特征损失,避免了模型退化和过拟合现象的发生;

(5)在 DRSN 模型中引入软阈值化和注意力机制,针对不同等级的含噪信号,自动设置阈值,提高了模型的抗干扰能力。

在后续的研究中,为了进一步提高模型的故障诊断正确率,增强其鲁棒性,笔者将选择更多实际工况下的样本数据对模型进行训练,并探索在更多深度学习模型中,引入残差收缩模块的可能性。

参考文献(References):

[1] 崔 灿,陈林波,张 敏. CNN 混合模型的旋转机械故障诊断方法研究[J]. 现代电子技术,2022,45(13):73-77.

- [2] EREN L, INCE T, KIRANYAZ S. A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier [J]. **Journal of Signal Processing Systems**,2019,91(2):179-189.
- [3] GLOWACZ A, GLOWACZ W, GLOWACZ Z, et al. Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals [J]. **Measurement**, 2017,113(1):1-9.
- [4] SONG L Y, WANG H Q, CHEN P. Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery [J]. **IEEE. Transactions on Instrumentation and Measurement**,2018,67(8):1887-1899.
- [5] LU Q, YANG R, ZHONG M, et al. An improved fault diagnosis method of rotating machinery using sensitive features and RLS-BP neural network [J]. **IEEE. Transactions on Instrumentation and Measurement**, 2020,69(4):1585-1593.
- [6] 李 政,张 炜,明安波,等. 基于 IEWT 和 MCKD 的滚动轴承故障诊断方法[J]. 机械工程学报,2019,55(23):136-146.
- [7] WANG F, DENG G, MA L, et al. Convolutional neural network based on spiral arrangement of features and its application in bearing fault diagnosis [J]. **IEEE. Access**, 2019,7:64092-64100.
- [8] KHODJA A Y, N GUERSI, SAADI M N, et al. Rolling element bearing fault diagnosis for rotating machinery using vibration spectrum imaging and convolutional neural networks [J]. **The International Journal of Advanced Manufacturing Technology**,2019,106(5-6):1737-1751.
- [9] WANG H, XU J, YAN R, et al. A new intelligent bearing fault diagnosis method using SDP representation and SE-CNN [J]. **IEEE. Transactions on Instrumentation and Measurement**,2020,69(5):2377-2389.
- [10] 余传粮,梁睿君,冉文丰,等. 基于 STFT 和 CNN 的齿轮箱故障诊断[J]. 机械制造与自动化,2022,51(3):152-154,195.
- [11] 高 晗,毛 阆,韦荣阳,等. 基于模型采样和领域对抗神经网络的轴承故障诊断方法[J]. 机电工程,2022,39(8):1024-1030.
- [12] HU Zong-xu, WANG Yan, GE Ming-feng, et al. Data-driven fault diagnosis method based on compressed sensing and improved multi-scale network [J]. **IEEE. Transactions on Industrial Electronics**, 2020, 67(4):3216-3225.
- [13] DUAN Li-xiang, XIE Meng-yun, WANG Jin-jiang, et al. Deep learning enabled intelligent fault diagnosis: overview and applications [J]. **Journal of Intelligent and Fuzzy Systems**,2018,35(3):1-14.

- [14] CHEN Xue-jun, YANG Yong-ming, CUI Zhi-xin, et al. Vibration fault diagnosis of wind turbine based on variational mode decomposition and energy entropy [J]. **Energy**, 2019, 174(5):1100-1109.
- [15] 车畅畅, 王华伟, 倪晓梅, 等. 基于深度学习的航空发动机故障融合诊断[J]. 北京航空航天大学学报, 2018, 44(3):621-628.
- [16] 许旭, 陈志刚, 杜小磊, 等. 基于同步挤压 S 变换和 DRN 的压裂车故障诊断研究[J]. 机电工程, 2019, 36(12):1325-1330.
- [17] 许启发, 程启亮, 蒋翠侠, 等. 基于组序列多分支 CNN-LSTM 的风机轴承和齿轮箱故障诊断研究[J]. 机电工程, 2022, 39(8):1050-1060.
- [18] PENG D, WANG H, LIU Z, et al. Multibranch and multiscale CNN for fault diagnosis of wheelset bearings under strong noise and variable load condition[J]. **IEEE Transactions on Industrial Informatics**, 2020, 16(7):4949-4960.
- [19] WEN L, LI X, GAO L, et al. A new convolutional neural network-based data-driven fault diagnosis method [J]. **IEEE Transactions on Industrial Electronics**, 2018, 65(7):5990-5998.
- [20] 温竹鹏, 陈捷, 刘连华, 等. 基于小波变换和优化 CNN 的风电齿轮箱故障诊断[J]. 浙江大学学报:工学版, 2022, 56(6):1212-1219.
- [21] LIANG P F, DENG C, WU J, et al. Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial net and convolutional neural network [J]. **Measurement**, 2020, 159(7):107768.
- [22] 卢锦玲, 郭鲁豫. 基于改进深度残差收缩网络的电力系统暂态稳定评估[J]. 电工技术学报, 2021, 36(11):2233-2244.
- [23] 辛阔, 王建国, 张文兴. 基于深度 Q 学习策略的旋转机械故障诊断方法研究[J]. 机电工程, 2021, 38(10):1261-1268.
- [24] 丁显, 徐进, 滕伟, 等. 基于无参数经验小波变换的风电齿轮箱故障特征提取[J]. 振动与冲击, 2020, 39(8):99-105.
- [25] ZHOU Quan, LI Yi-bing, TIAN Yu, et al. A novel method based on nonlinear auto-regression neural network and convolutional neural network for imbalanced fault diagnosis of rotating machinery [J]. **Measurement**, 2020, 161(9):107-180.
- [26] LI Xiang, ZHANG Wei, DING Qian. Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction [J]. **Reliability Engineering & System Safety**, 2019, 182(2):208-218.
- [27] WANG S X, CHEN H W. A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network [J]. **Applied Energy**, 2019, 235(2):1126-1140.
- [28] HAN Y, TANG B P, DENG L. An enhanced convolutional neural network with enlarged receptive fields for fault diagnosis of planetary gearboxes [J]. **Computers in Industry**, 2019, 107(5):50-58.
- [29] 车畅畅, 王华伟, 倪晓梅, 等. 基于深度残差收缩网络的滚动轴承故障诊断[J]. 北京航空航天大学学报, 2021, 47(7):1399-1405.
- [30] 樊显涛, 唐博文, 谭蔚. 并排塔器流致振动特性及预测模型研究[J]. 压力容器, 2022, 39(2):1-10.
- [31] WANG B, LEI Y, LI N, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings [J]. **IEEE Transactions on Reliability**, 2020, 69(1):401-412.

[编辑: 司红岩]

本文引用格式:

杨正理, 吴馥云, 陈海霞. 基于改进深度残差收缩网络的旋转机械故障诊断[J]. 机电工程, 2023, 40(3):344-352.

YANG Zheng-li, WU Fu-yun, CHEN Hai-xia. Fault diagnosis of rotating mechanical based on improved deep residual shrinkage network [J]. *Journal of Mechanical & Electrical Engineering*, 2023, 40(3):344-352.

《机电工程》杂志: <http://www.meem.com.cn>