

基于神经常微分方程的机械故障诊断方法

郭彪¹, 李志农^{1,2}

(1. 南昌航空大学无损检测技术教育部重点实验室, 江西 南昌 330063;

2. 汕头大学智能制造技术教育部重点实验室, 广东 汕头 515063)

摘要: 针对传统的深度学习故障诊断方法中存在架构可解释性差以及盲目堆叠层数导致的参数增加和内存消耗等问题, 将神经常微分方程(neural ordinary differential equations, NODE)引入到机械故障诊断中。搭建面向机械故障诊断的神经常微分方程网络架构, 利用神经网络参数化隐藏状态的导数代替指定隐藏层的离散序列。通过构建故障数据与故障类型的非线性关系, 使用常微分方程求解器(ODE solver)完成对不同故障类别的分类任务, 形成一种端对端的故障诊断模式。将该方法应用到机械故障诊断领域, 搭建特定的神经常微分方程网络模型, 通过故障数据的输入完成对不同故障类别的分类任务。将该模型应用到航空发动机主轴轴承故障诊断中, 并与残差网络模型的故障诊断方法进行对比。试验结果表明, 在确保准确率不降低的情况下, 该方法不仅减少了内存消耗, 而且将模型参数数量减少了将近五倍。

关键词: 故障诊断; 神经常微分方程; 动力学系统; 残差网络

中图分类号: TH133.33 **文献标志码:** A **DOI:** 10.16385/j.cnki.issn.1004-4523.202307050

Mechanical fault diagnosis method based on neural ordinary differential equations

GUO Biao¹, LI Zhinong^{1,2}

(1. Key Laboratory of Nondestructive Testing Ministry of Education, Nanchang Hangkong University, Nanchang 330063, China; 2. Key Laboratory of Intelligent Manufacturing Technology (Shantou University), Ministry of Education, Shantou 515063, China)

Abstract: Based on the problems of poor interpretability, as well as parameter increase and memory consumption caused by blind stacking layers in traditional fault diagnosis method based on deep learning, Neural ordinary differential equation (NODE) is introduced into mechanical fault diagnosis, the network structure of NODE for machinery fault diagnosis is constructed. In the constructed structure, the derivatives of the parameterized hidden states of the neural network are used to replace the discrete sequences of the specified hidden layers. By constructing a nonlinear relationship between fault data and fault types, an ordinary differential equation solver (ODE solver) is used to complete the classification of different fault types, and an end-to-end fault diagnosis model is formed. The proposed method is applied to mechanical fault diagnosis to build a specific NODE network model, and the classification task of different fault categories is accomplished through the input of fault data. The constructed model is applied to the fault diagnosis of spindle bearing in the aircraft engine, and compared with the fault diagnosis method based on residual network model. The experimental results show that the constructed model and residual network model have satisfactory accuracy. However, the constructed model not only reduces the memory consumption, but also reduces the number of model parameters by almost five times.

Keywords: fault diagnosis; neural ordinary differential equation; dynamics system; residual network

在大数据背景下, 机械故障智能诊断的深入研究和应用迎来了新的机遇, 特别是基于深度学习的机械故障诊断方法取得了很大的进展。LIU 等^[1]利

用深度神经网络对旋转机械进行故障诊断, 并与贝叶斯、支持向量机等算法进行比较。JING 等^[2]利用深度卷积网络对齿轮箱进行自适应特征提取, 并证

收稿日期: 2023-07-17; **修订日期:** 2023-10-22

基金项目: 国家自然科学基金资助项目(52075236); 江西省自然科学基金重点项目(20212ACB202005); 智能制造技术教育部重点实验室(汕头大学)开放课基金资助项目(STME2024002); 广东省普通高校创新团队资助项目(2020KCXTD012)

明其相较于时、频域特征提取更有优势。ZHANG等^[3]将深度残差网络应用到滚动轴承故障诊断,有效防止了梯度爆炸和梯度消失,解决了网络退化的问题。时培明等^[4]在深度残差网络的基础上结合注意力机制,有效地提高了滚动轴承故障诊断的准确率。刘飞等^[5]为了克服现有基于深度学习的滚动轴承故障诊断算法训练参数量大、训练时间长且需要大量训练样本的缺点,提出了一种基于迁移学习(TL)与深度残差网络(residual network, ResNet)的快速故障诊断算法(TL-ResNet)。池福临等^[6]采用联合最大平均偏差准则和条件对抗学习域适配网络这两大机制实现了对齐源域和目标域,通过添加软阈值构建深度收缩残差网络提取噪声冗余下的轴承振动数据的特征信息。针对变负荷、变速度与变噪声三种试验设置验证了模型有效性。尚杰^[7]针对变转速工况下的滚动轴承故障问题,提出了一种结合同步提取变换和二维残差网络的故障诊断模型,该模型将时频图作为数据集,利用深度残差网络在数据集上进行训练,在噪声干扰下仍能保证诊断准确率。侯召国等^[8]提出了一种改进深度残差网络模型(improved deep residual network, IDRN),主要用来解决旋转机械标签样本不足而导致的故障特征提取困难问题。秦国浩等^[9]为了减小由于卷积核尺度单一导致的难以充分捕捉不同冲击的频率成分的问题,提出了一种动态宽卷积残差网络(dynamic wide-kernels residual network, DWResNet)轴承故障诊断方法。通过改进网络动态权重加权机制有效提高了在强背景噪声工况下轴承故障诊断的效果。

在这些基于数据驱动的机械故障智能诊断方法中,研究人员经常使用神经网络建立故障诊断模型。但是,随着人们对诊断精度的要求不断提高,使得神经网络的层数不断增加,结构不断深化。很多研究者开始不关注模型本身的计算成本,使得模型的内存成本和网络参数呈线性增长,导致模型难以被优化且易于过度拟合^[10],并且将模型应用到现实当中也会变得非常困难。为了增强深度学习模型的可解释性、减少层数堆叠造成的内存消耗且不降低诊断准确率,CHEN等^[11]提出残差网络^[12](ResNet)可以被解释为离散的常微分方程^[13](ordinary differential equations, ODE)的观点,合理地阐明了神经网络与动态系统两者之间的关系。其中,CHEN等^[11]在时间连续的前提下,利用常微分方程求解器对网络隐藏状态进行动力学建模,并使用被称为伴随灵敏度的方法^[11,14]代替机器学习中的反向传递算法。该模型在监督学习、时间序列预测和连续归一化流等问题的解决上表现出在误差控制、内存效率以及自适应计算等方面的优势。基于神经常微分方程的独特

优势,本文将神经常微分方程引入到机械故障诊断中,提出了基于神经常微分方程的机械故障诊断方法,为解决现有的智能诊断方法存在的架构可解释性差以及盲目堆叠层数导致的参数增加和内存消耗等问题提供了一种思路。为突出该方法的优越性,本文还将提出的方法与传统的残差网络模型的故障诊断方法进行了对比,试验结果验证了提出方法的有效性。

1 神经常微分方程

残差网络(ResNet)作为一种卷积神经网络,用于学习隐藏变量(如 x_{i+1} 和 x_i)之间的差值。ResNet的表达式如下:

$$x_{i+1} = x_i + f_i(x_i, \theta_i) \quad (1)$$

式中, x_i 为层数 $i=1, \dots, n$ 时的ResNet网络层的变量; f_i 为ResNet在第 i 层使用的神经网络函数; θ_i 为第 i 层的网络参数。

将神经常微分方程(NODE)作为ResNet的连续化引入^[11],并且把NODE表达式中网络层数 t 作为连续变量,无限细分,就可以写出其微分形式:

$$\frac{dx_t}{dt} = f(x_t, t, \theta) \quad (2)$$

式中, x_t 为层数 $t=1, \dots, n$ 时的变量; f 为由可微分的神经网络构成的微分方程; θ 为网络中可学习的参数^[15]。

若想要知道某层的隐藏状态,只需给出其初始状态 x_{t_0}, t_0, t_1 就可以利用下式进行积分计算:

$$x_{t_1} = x_{t_0} + \int_{t_0}^{t_1} f(x_t, t, \theta) dt \quad (3)$$

图1展示一个简单的神经常微分方程网络层传递的示例,可视化地说明了NODE如何获取输入变量并通过网络将其转换为输出变量的过程。图1中,每个实线箭头都是网络 $dx_t/dt=f(x_t, t)$ 的值,用于查找下一步网络的隐藏层变量,其中网络 f 为描述一阶ODE dx_t/dt 的函数, t 为网络的深度。使用数值微分方程求解器将输入通过神经网络 f 描述

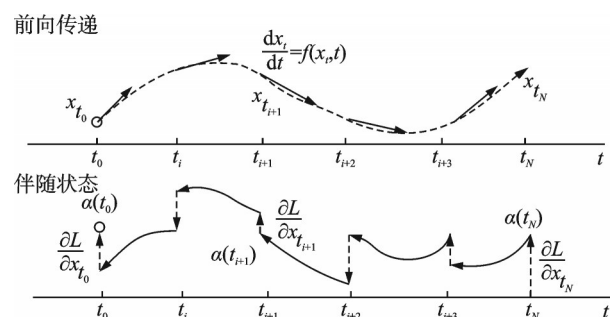


图1 神经常微分方程前向传递和伴随灵敏度方法
Fig. 1 Forward and the adjoint sensitivity method of NODE

的向量场,并最终转变为输出变量。在这个过程中,依靠在常微分方程求解器中选择步长,可以控制每一层函数计算的数量。图1中五个箭头相当于具有五个隐藏层的网络。标记为 x_t 的虚线路径是变量通过NODE的轨迹, x_{t_n} 为NODE的最终输出。由于神经网络的参数 θ 不随变量在向量场中的传递而改变,为了简便,网络将使用 $f(x_t, t)$ 取代 $f(x_t, t, \theta)$ 。

以上为NODE的前向传递过程。为了在反向传递中节省内存成本,并且避免引入额外的数值计算误差,采用伴随灵敏度方法,舍弃传统的反向传递算法。其核心思想是采用伴随灵敏度方法取代传统的梯度传递,而不再通过前向传递中的常微分方程求解器。具体过程如下,若将前向计算的损失函数记为 L ,则整个常微分方程求解器的前向损失为:

$$L(x_{t_n}) = L\left(x_{t_0} + \int_{t_0}^{t_n} f(x_t, t, \theta) dt\right) \quad (4)$$

其伴随状态的定义为:

$$\alpha(t) = \frac{\partial L}{\partial x_t} \quad (5)$$

式中, $\alpha(t)$ 表示在变量 t 下的关于隐藏状态 x 的伴随状态。

在NODE反向传递的过程中使用伴随灵敏度法, $\alpha(t)$ 为损失 L 对 x_t 的导数。可以在连续时间下通过下式将梯度继续向前传递:

$$\frac{d\alpha(t)}{dt} = -\alpha(t) \frac{\partial f(x_t, t, \theta)}{\partial x} \quad (6)$$

$$\alpha(t_0) = \alpha(t_1) - \int_{t_1}^{t_0} \alpha(t) \frac{\partial f(x_t, t, \theta)}{\partial x_t} dt \quad (7)$$

将上式从 t_1 伴随状态到 t_0 伴随状态的计算过程拓展至整个反向传递网络,计算此过程的整个梯度:

$$\frac{dL}{d\theta} = \int_{t_1}^{t_0} \alpha(t) \frac{\partial f(x_t, t, \theta)}{\partial \theta} dt \quad (8)$$

以上为NODE使用伴随灵敏度法完成反向传递的过程,由于此方法不需要保存反向传递过程中产生的导数梯度,只需记录最后一个时间步的隐藏状态,而神经网络需保存导数梯度、权重和偏置等参数来完成导数链式法则的传递,故减少了不必要的内存消耗。该方法之所以可以替代神经网络的反向传递,是在前向传递中使用了神经微分方程,计算出了每层的微分形式表示的隐藏状态。而且选择不同的常微分方程求解方法,更有利于平衡计算时间和准确率。

基于此,许多神经网络可以被微分方程所替换,利用常微分方程求解器取代各神经元及各层之间复杂的计算过程,避免盲目堆叠层数导致的浪费,可以有效减少参数数量,实现参数共享,显著减少内存成本^[16]。

2 神经微分方程故障诊断模型的建立

由于神经微分方程可以被视为连续化的残差网络,为了更好地说明其在故障诊断方面的优势,分别建立残差网络和神经微分方程网络两个模型进行对比。残差网络(ResNet)模型共有三层结构,分别为下采样卷积层、残差网络层和前馈层。如图2所示,其下采样层用来接收输入数据,进而通过设置残差块(ResBlock)中第一个Conv2D层的步幅来提取特征并降低数据的维数。下采样层的输出作为第一个ResBlock的输入,之后依次传递给下一个ResBlock。像这样连接 n 个ResBlock,最终ResBlock的输出被展平并通过前馈网络,该网络将输出七种分类标签。表1为所用的ResNet架构的基本布局以及每一层输出张量的大小。表2指出架构中使用的ResBlock的构成。ResNet模型的超参数选择如表3所示。

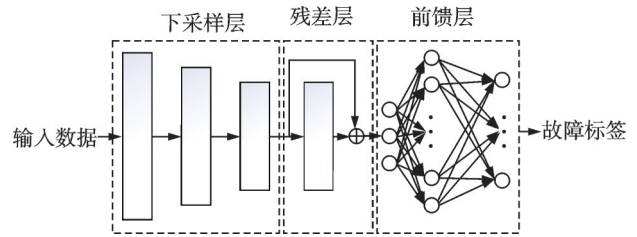


图2 残差网络模型框架图

Fig. 2 ResNet model framework diagram

表1 残差网络的模型架构及每层输出张量

Tab. 1 Model architecture and output tensor per layer of ResNet

网络层	层架构	输出
	Conv2D	(64, 64, 1020, 1)
下采样层	ResBlock	(64, 64, 510, 1)
	ResBlock	(64, 64, 255, 1)
残差层	ResBlock × 6	(64, 64, 255, 1)
	GroupNorm	(64, 64, 255, 1)
	ReLU	(64, 64, 255, 1)
前馈层	Adaptive AvgPool2D	(64, 64, 1, 1)
	Flatten	(64, 64)
	Linear	(64, 7)

表2 残差块的构成

Tab. 2 Composition of ResBlock

层架构	输出
GroupNorm	(64, 64, 255, 1)
ReLU	(64, 64, 255, 1)
Conv2D	(64, 64, 255, 1)
GroupNorm	(64, 64, 255, 1)
Conv2D	(64, 64, 255, 1)

表 3 残差网络模型超参数选择

Tab. 3 Hyperparameters selection for ResNet model

超参数	取值
Epochs	160
Train_Batch_size	64
Test_Batch_size	128
Optimizer_SGD_momentum	0.9
Learning rate	0.1
Amount of ResBlock(n)	6

与残差网络模型(图 2)相似,神经常微分方程(NODE)模型(图 3)也共有三层结构,包括下采样层、神经常微分方程层和前馈层。下采样层同样用来提取特征并降低输入数据的维数。两种模型结构的主要区别在于中间层,使用常微分方程块(ODE-block)代替 ResBlock,其输出通过数值常微分方程求解器向前传递轨迹,将离散化的网络变为连续,以

微分方程的形式改进了神经网络各层之间的传递方式(t 在此过程中不是时间,而是节点网络的深度, $f(x, t)$ 表示任意一点处的梯度)。之后将隐藏状态输入到前馈层,通过前馈网络的展平层和线性层,最后以七种分类标签的形式输出。表 4 展示了 NODE 架构的基本布局以及每一层输出张量的大小。NODE 模型的超参数选择如表 5 所示。

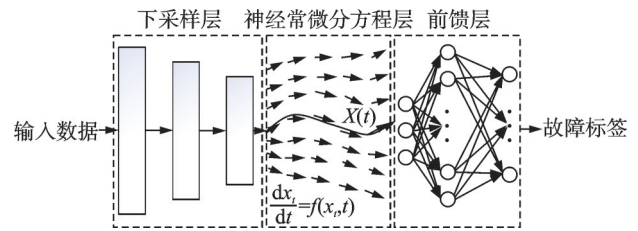


图 3 神经常微分方程网络模型框架图
Fig. 3 NODE model framework diagram

表 4 神经常微分方程网络模型架构及每层输出张量

Tab. 4 Model architecture and output tensor per layer of NODE

网络层	层架构	输出
下采样层	Conv2D	(64, 64, 1020, 1)
	GroupNorm	(64, 64, 1020, 1)
	ReLU	(64, 64, 1020, 1)
	Conv2D	(64, 64, 510, 1)
	GroupNorm	(64, 64, 510, 1)
	ReLU	(64, 64, 510, 1)
	Conv2D	(64, 64, 255, 1)
	GroupNorm	(64, 64, 255, 1)
	ReLU	(64, 64, 255, 1)
神经常微分方程层	Conv2D	(64, 64, 255, 1)
	GroupNorm	(64, 64, 255, 1)
	Conv2D	(64, 64, 255, 1)
	GroupNorm	(64, 64, 255, 1)
	GroupNorm	(64, 64, 255, 1)
	ReLU	(64, 64, 255, 1)
	Adaptive AvgPool2D	(64, 64, 1, 1)
	Flatten	(64, 64)
	Linear	(64, 7)

表 5 神经常微分方程网络模型超参数选择

Tab. 5 Hyperparameters selection for NODE model

超参数	取值
Epochs	160
Train_Batch_size	64
Test_Batch_size	128
Optimizer_SGD_momentum	0.9
Learning rate	0.1
ODE solver	RK4

3 试验研究

试验采用都灵理工大学航空主轴轴承数据^[17]进行故障诊断。试验台如图 4 所示,具体故障数据集

如表 6 所示。选择在负载 1000 N,转速 200 Hz 条件下 x 方向上采集的数据。7 类工况包括正常工况,故障维度分别为 150、250 和 450 μm 下的轴承内圈故障以及故障维度分别为 150、250 和 450 μm 下的滚动体故障。对每种工况采集的振动信号划分样本,每类工况构造 400 个样本,共 2800 个样本,每个样本选取 1024 个采样点。将总样本数按照 4:1 的比例划分训练集和测试集。

图 5 为 5 种 ODE solver: RK4 (Runge-Kutta), Bosh3(Bogacki-Shampine), Dopri5(Dormand-Prince), Dopri8^[18]和 Euler 分别训练 NODE 模型所用时间。图 6 为五种 ODE solver 训练的函数评估次数(number of function evaluations, NFE), NFE 越小代表其计算效率与内存效率越高。从图中可以看出, Euler 方法

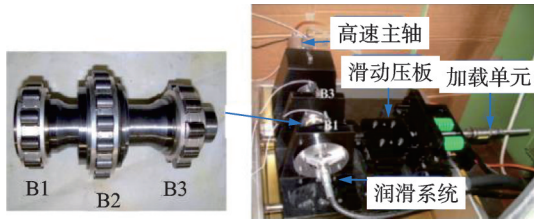


图4 轴承试验台

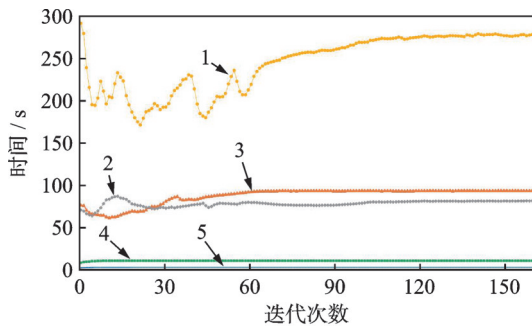
Fig. 4 Test rig of bearing

表6 航空发动机主轴轴承数据集

Tab. 6 Aero-engine spindle bearing dataset

健康状态	故障维度/ μm	额定转速/ Hz	额定负载/ N	方向
正常/内圈	150/250/	100/200/	0/1000/	x/y/z
故障/滚动	450	300/400/500	1400/1800	
体故障				

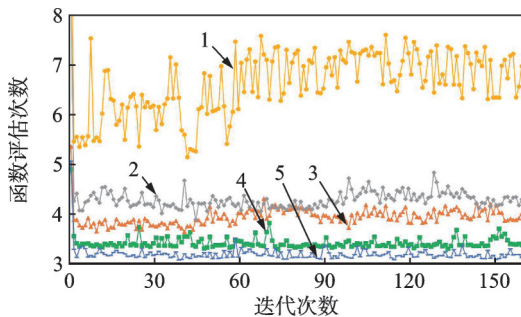
训练所用时间和 NFE 最少, RK4 次之。但由于 Euler 法为一阶数值方法, 当步数增多时, 误差会因积累而越来越大。RK4 为四阶方法, 其精度相较于 Euler 法有很大程度的提高。所以本次试验 NODE 故障诊断模型中常微分方程求解器选 RK4。



1: Dopri8; 2: Dopri5; 3: Bosh3; 4: RK4; 5: Euler

图5 训练所用时间

Fig. 5 Training time

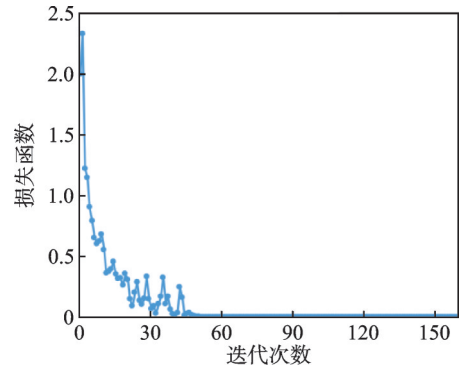


1: Dopri8; 2: Dopri5; 3: Bosh3; 4: RK4; 5: Euler

图6 函数评估次数

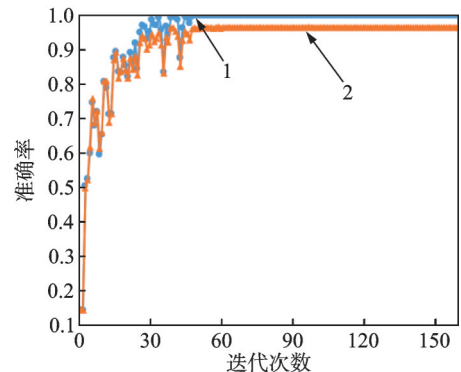
Fig. 6 Number of function evaluations

ResNet 模型和 NODE 模型训练和测试的结果分别如图 7 和 8 所示。图 7(a) 和 8(a) 分别展示了两种模型训练集上的损失函数曲线。可以看出损失函数值随着迭代次数的增加都逐渐收敛并趋向



(a) 损失函数曲线

(a) The curve of loss function



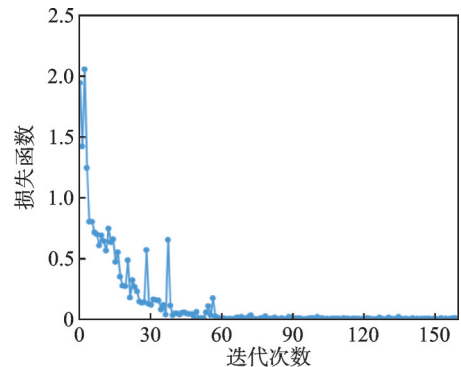
1: 训练集准确率; 2: 测试集准确率

(b) 准确率曲线

(b) The curve of accuracy

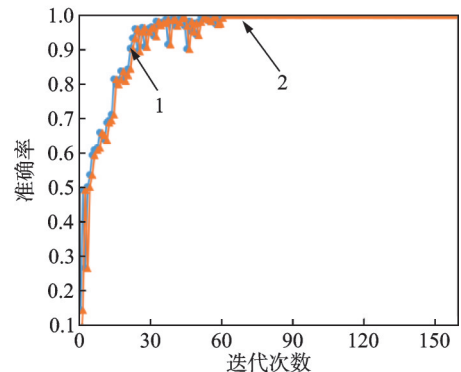
图7 残差网络模型诊断曲线

Fig. 7 Diagnostic curve of ResNet model



(a) 损失函数曲线

(a) The curve of loss function



1: 训练集准确率; 2: 测试集准确率

(b) 准确率曲线

(b) The curve of accuracy

图8 神经常微分方程网络模型诊断曲线

Fig. 8 Diagnostic curve of NODE model

于 0,反映了模型找到了最优值,并没有陷入过拟合状态^[19]。图 7(b)和 8(b)展示了两种模型的训练集和测试集的准确率曲线,在训练集上两种模型的准确率都能达到 100%。ResNet 模型和 NODE 模型测试集的最终准确率分别为 96.61% 和 98.93%。

表 7 给出了两种模型在准确率、参数数量和内存占用的对比。

表 7 两种模型试验结果对比

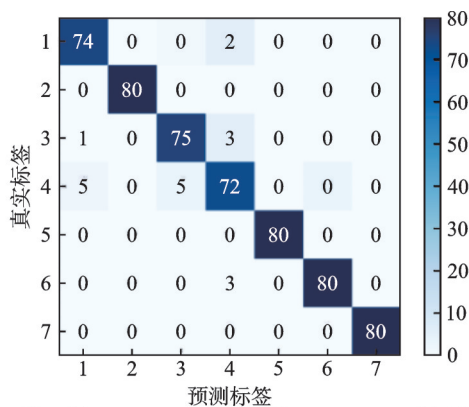
Tab.7 Comparison of experimental results of the two models

模型指标	ResNet	NODE
准确率/%	96.61	98.93
参数数量	601031	117703
内存占用/MiB	3356	2654

从表 7 中可以看出,NODE 模型较 ResNet 模型在测试集上准确率略有提高,而且通过 pytorch 框架自带函数,计算 ResNet 模型所需的参数数量(即各层之间参数+每层的偏差)几乎为 NODE 模型的 6 倍,同时 NODE 模型的内存占用更少,这表明了 NODE 通过使用常微分方程求解器代替神经网络起到了优化模型的作用。

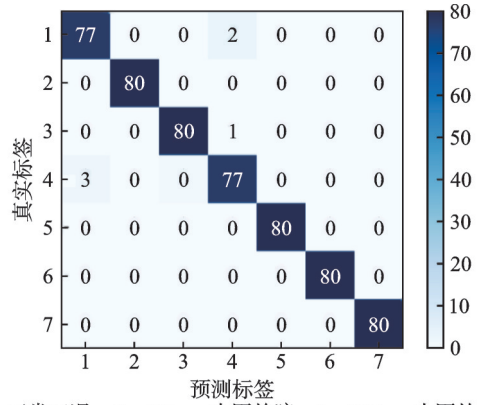
图 9 为 ResNet 故障诊断模型在测试集上的混淆矩阵,共 560 个样本,误判 19 个样本。其中 1 个真实标签为故障维度 250 μm 的内圈故障和 5 个真实标签为故障维度 150 μm 的内圈故障被误判为正常; 5 个真实标签为故障维度 150 μm 的内圈故障被误判为故障维度 250 μm 的内圈故障; 2 个真实标签为正常、3 个故障维度 250 μm 的内圈故障和 3 个故障维度 250 μm 的滚动体故障被误判为故障维度 150 μm 的内圈故障。

图 10 为 NODE 故障诊断模型在测试集上的混淆矩阵,共 560 个样本,误判 6 个样本。其中 3 个真



1: 正常工况; 2: 450 μm 内圈故障; 3: 250 μm 内圈故障; 4: 150 μm 内圈故障; 5: 450 μm 滚动体故障; 6: 250 μm 滚动体故障; 7: 150 μm 滚动体故障

图 9 残差网络模型的混淆矩阵
Fig.9 Confusion matrix for ResNet models



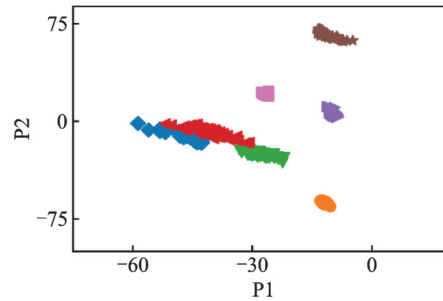
1: 正常工况; 2: 450 μm 内圈故障; 3: 250 μm 内圈故障; 4: 150 μm 内圈故障; 5: 450 μm 滚动体故障; 6: 250 μm 滚动体故障; 7: 150 μm 滚动体故障

图 10 神经常微分方程网络模型的混淆矩阵
Fig.10 Confusion matrix for NODE models

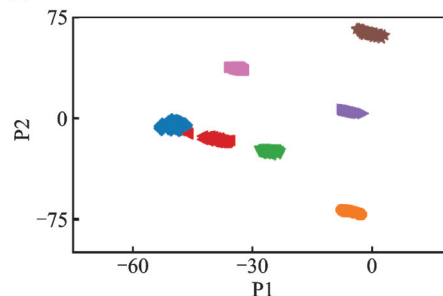
实标签为故障维度 150 μm 的内圈故障被误判为正常; 2 个真实标签为正常和 1 个真实标签为故障维度 250 μm 的内圈故障被误判为 150 μm 的内圈故障。

为了进一步观察 NODE 故障诊断模型的有效性,利用 t-分布邻域嵌入(t-distribution stochastic neighbor embedding, t-SNE^[20])算法对输出的高维故障特征进行降维,并以散点图的形式在二维平面上进行展示(散点图的坐标轴无实际意义,用 P1 和 P2 表示)。使用 ResNet 模型对测试集进行分类,结果如图 11(a)所示。其中正常、故障维度 150 μm 和

- — 450 μm 内圈故障
- ▲ — 450 μm 滚动体故障
- ◆ — 正常工况
- ▼ — 250 μm 内圈故障
- ★ — 250 μm 滚动体故障
- ◀ — 150 μm 内圈故障
- — 150 μm 滚动体故障



(a) 残差网络模型分类结果的可视化
(a) Visualization of classification results for ResNet models



(b) 神经常微分方程网络模型分类结果的可视化
(b) Visualization of classification results for NODE models

图 11 ResNet 和 NODE 模型在测试集上的分类结果可视化
Fig.11 Visualization of classification results of ResNet and NODE models on the test dataset

250 μm 的内圈故障这三个类别之间存在明显的特征重叠,其余四种类别的故障特征之间相互分离。使用 NODE 模型对测试集进行分类,结果如图 11 (b)所示,图中七种轴承工况的特征已经聚类明显,除了正常和故障维度 150 μm 的内圈故障这两个类别中个别存在特征重叠外,其余五种类别故障特征之间相互分离。这也呼应了图 10 混淆矩阵中正常和故障维度 150 μm 的内圈故障这两种类别之间存在的误判,直观反映了 NODE 故障诊断模型相较于 ResNet 可以更好地对轴承不同工况进行有效识别。

4 结 论

本文从神经网络与动力学系统两者相互联系的角度入手,阐明了可以将神经网络视作常微分方程的观点,使得神经网络中的黑盒模型具备了可解释性。之后为了提高模型的准确率,减少模型参数,将残差网络的中间层替换为常微分方程层,构建了神经常微分方程网络模型。并将该模型引入到机械故障诊断领域。以都灵理工大学航空主轴轴承数据作为试验数据,通过端对端的故障诊断模式,完成了七种工况下轴承故障分类的任务。与残差网络模型进行对比,在测试集上提高了故障识别准确率的情况下,减少了内存成本并将模型参数减小至残差网络模型的六分之一,证明了通过使用常微分方程代替神经网络起到了优化模型的作用。

参考文献:

- [1] LIU R N, YANG B Y, ZIO E, et al. Artificial intelligence for fault diagnosis of rotating machinery: a review [J]. *Mechanical Systems and Signal Processing*, 2018, 108: 33-47.
- [2] JING L Y, ZHAO M, LI P, et al. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox[J]. *Measurement*, 2017, 111: 1-10.
- [3] ZHANG W, LI X, DING Q. Deep residual learning-based fault diagnosis method for rotating machinery[J]. *ISA Transactions*, 2019, 95: 295-305.
- [4] 时培明,吴术平,于越,等.基于注意力机制和深度残差网络的滚动轴承故障诊断[J].*燕山大学学报*,2024,48(1): 39-47.
SHI Peiming, WU Shuping, YU Yue, et al. Rolling bearing fault diagnosis based on attention mechanism and depth residual network[J]. *Journal of Yanshan University*, 2024, 48(1): 39-47.
- [5] 刘飞,陈仁文,邢凯玲,等.基于迁移学习与深度残差网络的滚动轴承快速故障诊断算法[J].*振动与冲击*, 2022, 41(3): 154-164.
LIU Fei, CHEN Renwen, XING Kailing, et al. Fast fault diagnosis algorithm for rolling bearing based on transfer learning and deep residual network[J]. *Journal of Vibration and Shock*, 2022, 41(3): 154-164.
- [6] 池福临,杨新宇,邵思羽,等.基于深度收缩残差网络的轴承变工况故障诊断[J].*计算机集成制造系统*, 2023, 29(4): 1146-1156.
CHI Fulin, YANG Xinyu, SHAO Siyu, et al. Bearing fault diagnosis under variable working conditions based on deep residual shrinkage networks[J].*Computer Integrated Manufacturing Systems*, 2023, 29(4): 1146-1156.
- [7] 尚杰.基于深度残差网络的旋转机械故障诊断[D].青岛:青岛理工大学,2021.
SHANG Jie. Fault diagnosis of rotating machinery based on deep residual network[D]. Qingdao: Qingdao University of Technology, 2021.
- [8] 侯召国,王华伟,周良,等.基于改进深度残差网络的旋转机械故障诊断[J].*系统工程与电子技术*, 2022, 44(6): 2051-2059.
HOU Zhaoguo, WANG Huawei, ZHOU Liang, et al. Fault diagnosis of rotating machinery based on improved deep residual network[J]. *Systems Engineering and Electronics*, 2022, 44(6): 2051-2059.
- [9] 秦国浩,张楷,丁昆,等.动态宽卷积残差网络的轴承故障诊断方法[J].*中国机械工程*, 2023, 34(18): 2212-2221.
QIN Guohao, ZHANG Kai, DING Kun, et al. Dynamic wide convolutional residual network for bearing fault diagnosis method[J]. *China Mechanical Engineering*, 2023, 34(18): 2212-2221.
- [10] 李亮.基于神经常微分方程的时序数据分析与预测[D].成都:电子科技大学,2021.
LI Liang. Analyzing and forecasting time-series data via neural ordinary differential equations[D]. Chengdu: University of Electronic Science and Technology of China, 2021.
- [11] CHEN R T Q, RUBANOVA Y, BETTENCOURT J, et al. Neural ordinary differential equations[C]// *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. 2018: 6572-6583.
- [12] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 770-778.

- [13] WEINAN E. A proposal on machine learning via dynamical systems[J]. *Communications in Mathematics and Statistics*, 2017, 5(1): 1-11.
- [14] DUPONT E, DOUCET A, TEH Y W. Augmented neural ODEs[C]// *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. 2019: 3140-3150.
- [15] NORCLIFFE A, BODNAR C, DAY B, et al. On second order behaviour in augmented neural ODEs[J]. *Advances in Neural Information Processing Systems*, 2020, 33: 5911-5921.
- [16] KIDGER P. On neural differential equations[J]. *arXiv preprint arXiv:2202.02435*, 2022.
- [17] DAGA A P, FASANA A, MARCHESIELLO S, et al. The Politecnico di Torino rolling bearing test rig: description and analysis of open access data[J]. *Mechanical Systems and Signal Processing*, 2019, 120: 252-273.
- [18] HAIRER E, NORSETT S P, WANNER G. *Solving Ordinary Differential Equations I : Nonstiff Problems* [M]. 2nd ed. Springer-Verlag, 1993.
- [19] 李泽东, 李志农, 王成军. 深度卷积神经网络在轴承多故障复合诊断中应用研究[J]. *南昌航空大学学报(自然科学版)*, 2020, 34(1): 12-20.
LI Zedong, LI Zhinong, WANG Chengjun. Application of deep convolution neural network in bearing multi fault diagnosis[J]. *Journal of Nanchang Hangkong University (Natural Sciences)*, 2020, 34(1): 12-20.
- [20] VAN DER MAATEN L, HINTON G. Visualizing Data using t-SNE[J]. *Journal of Machine Learning Research*, 2008, 9: 2579-2605.

第一作者: 郭 彪(1998—),男,硕士研究生。

E-mail: guobiao980315@163.com

通信作者: 李志农(1966—),男,博士,教授。

E-mail: lizhinong@tsinghua.org.cn