DOI:10.19344/j.cnki.issn1671-5276.2024.06.048

残差胶囊网络在旋转机械故障诊断中的应用研究

吴萍,李曙生

(泰州职业技术学院,江苏 泰州 225316)

摘　要:针对旋转机械中的故障诊断需求,在传统的胶囊网络中引入残差块和模糊 C 均值聚类算法,构建残差胶囊网络故障诊断模型。在残差胶囊网络的基础上,引入注意力机制和 G-K 动态路由算法,构建注意力胶囊网络故障诊断模型。仿真分析表明:两种模型都能对故障进行精准测试,具有较强的表达能力和泛化能力。

关键词:旋转机械;故障诊断;胶囊网络;残差块;注意力机制

中图分类号:TP277　　**文献标志码**:B　　**文章编号**:1671-5276(2024)06-0244-03

Research on Application of Residual Capsule Network in Fault Diagnosis of Rotating Machinery

WU Ping, LI Shusheng

(Taizhou Polytechnic College, Taizhou 225316, China)

Abstract: To meet the fault diagnosis requirements in rotating machinery, a residual capsule network fault diagnosis model is constructed by introducing residual blocks and fuzzy C-means clustering algorithm in traditional capsule networks. On the basis of residual capsule network, attention mechanism and G-K Dynamic routing algorithm are introduced to build a fault diagnosis model of attention capsule network. Simulation analysis shows that both models can accurately test faults and have strong expressive and generalization abilities.

Keywords: rotating machinery; fault diagnosis; capsule network; residual block; attention mechanism

0　引言

在工业 4.0 时代的背景下,机械故障诊断的核心目标是利用先进的技术手段提高机械设备使用寿命,减少由于机械故障带来的经济损失[1]。然而机械故障涉及的机制较复杂,不能通过简单的实验和检测来完全捕获每个故障的特征,而现有的诊断模型往往只能够诊断单个机械部件的故障。在实际的应用中,故障的产生是由多个部件共同导致的。基于此,本文在传统的胶囊网络中引入残差块和模糊 C 均值聚类算法构建残差胶囊网络故障诊断模型,并在残差胶囊网络的基础上,引入注意力机制和 G-K 动态路由算法,构建注意力胶囊网络故障诊断模型,以提升旋转机械模型在故障诊断中的准确率和泛化能力。

1　基于残差胶囊网络的旋转机械故障诊断模型构建

1.1　残差胶囊网络故障诊断模型构建

残差神经网络(residual neural network, ResNet)在深度神经网络模型中引入了残差块来解决深度网络中的梯度消失和梯度爆炸问题[2]。在残差块中,输入的特征与输出的特征通过跨层连接相加,在保留原始特征信息的基础上,能够避免梯度消失和梯度爆炸。损失函数对残差块输入值的导数如式(1)所示。

$$\frac{\partial E}{\partial X_l}=\frac{\partial E}{\partial X_{l+1}}\left(1+\frac{\partial F(X_l)}{\partial X_l}\right) \tag{1}$$

式中:E 表示损失函数;X_{l+1}表示残差块的输入值;$F(X_l)$表示残差函数;X_l 表示输入值的映射。通过残差函数对输入值进行求导后,当导数为 0 时,$\left(1+\frac{\partial F(X_l)}{\partial X_l}\right)$值最小为 1,避免了梯度消失的问题。引入模糊 C 均值聚类算法对动态路由进行改进,并将隶属度函数和聚类中心更新,以达到最优解的目的[3]。隶属度矩阵如式(2)所示。

$$\boldsymbol{w}_{ij}=\frac{1}{\sum_{k=1}^{l}\left(\frac{\|u_i-v_j\|}{\|u_i-v_k\|}\right)^{\left(\frac{2}{m-1}\right)}} \tag{2}$$

式中:u_i 表示第 i 个样本;v_j 表示第 j 个簇的聚类

第一作者简介:吴萍(1969—),女,江苏泰兴人,高级工程师,高级实验师,本科,研究方向为现代机械制造技术,905110769@qq.com。

中心;l 表示聚类中心个数;m 表示模糊程度;$\|u_i-v_k\|$表示两者间的欧式距离。聚类中心的计算公式如式(3)所示。

$$v_{ij}=\frac{\sum_{i=1}^{n}\boldsymbol{w}_{ij}^{m}u_i}{\sum_{i=1}^{n}\boldsymbol{w}_{ij}^{m}} \tag{3}$$

式中 $\boldsymbol{w}_{ij}^{m}$表示更新后的隶属度矩阵。将胶囊网络和残差块相结合,构建残差胶囊网络(residual capsule network, RCN)来解决滚动轴承中的故障诊断问题[4]。RCN 结构如图 1 所示。

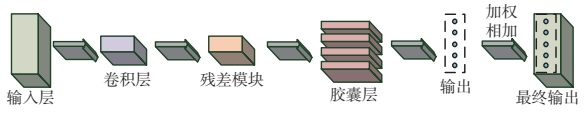


图 1　RCN 结构示意图

胶囊层的输出包含一个实体的存在概率和属性向量,与传统的卷积神经网络不同,RCN 中的残差块不仅将输入向量加到输出向量上,还能以一定的方式重新加权相加。采用边际损失函数作为损失函数,公式如式(4)所示。

$$L_k=T_k\max(0,m^{+}-\|v_k\|)^2+\lambda(1-T_k)\max(0,\|v_k\|-m^{-})^2 \tag{4}$$

式中:k 表示分类类别;T_k 表示指标函数;m^{+}表示上边界;m^{-}表示下边界;λ 表示比例系数。残差胶囊网络结合了残差块和胶囊网络的优点,通过残差连接和动态路由算法保持了一定的梯度流,避免了梯度消失和爆炸的问题。

1.2　基于注意力胶囊网络的旋转机械故障诊断模型构建

残差注意力模块(residual attention module, RAM)是一种结合了注意力机制和残差连接的神经网络模块,其中两个卷积层和残差模块构成了主分支,注意力分支用于特征动态权重计算[5]。注意力模块的输出公式如式(5)所示。

$$H_{i,c}(x)=M_{i,c}(x)\times T_{i,c}(x) \tag{5}$$

式中:$M_{i,c}(x)$表示掩膜网络的输出;$T_{i,c}(x)$表示主干网络的输出。采用 G-K 模糊聚类算法对生成的特征进行聚类操作,公式如式(6)所示。

$$J(X,\boldsymbol{U},V)=\sum_{i=1}^{c}\sum_{k=1}^{N}(\boldsymbol{u}_{ik})m\,\|x-v_i\|_{M_i}^{2} \tag{6}$$

式中:v_i 表示聚类中心;$\boldsymbol{u}_{ik}$表示隶属度矩阵;X 表示数据集;$\boldsymbol{U}$ 表示模糊矩阵;V 表示聚类中心集合;c 表示模糊类别;N 表示特征个数。针对旋转机械故障振动信号的特征,基于残差注意力模块,构建注意力胶囊网络故障诊断模型(attention myprotein capsule network, AMCN),结构如图 2 所示。

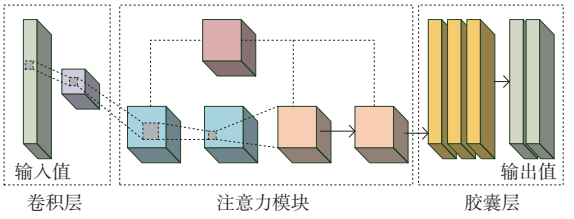


图 2　AMC 结构示意图

AMC 中包含卷积层、残差注意力模块和胶囊层。为了提高注意力胶囊网络中的真实特征提取概率,通过反向传播算法计算出损失函数对注意力权重和胶囊网络参数的梯度,并通过梯度下降等优化方式更新注意力权重和胶囊网络参数[6]。

2　故障诊断模型实例诊断及分析

2.1　残差胶囊网络模型轴承故障诊断及分析

为了测试残差胶囊网络在轴承故障诊断中的性能,实验选用 MFS-Magnum 试验台分别在 9.5 kHz、19.85 kHz、29.88 kHz 转频下对收集到的故障信息进行测试。数据特征从原始振动信号开始采样 50 个点,间隔 20 个点继续采样 6 次,总共 300 个测试数据。选择传统深度神经网络(deep neural network, DNN)、卷积神经网络(convolutional neural network, CNN)、ResNet 网络与 RCN 模型进行对比测试。不同模型的收敛曲线如图 3 所示。

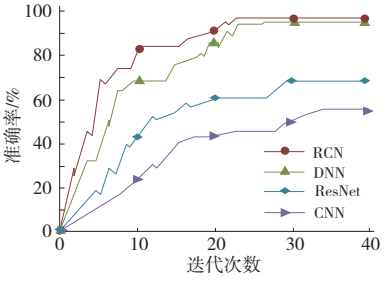
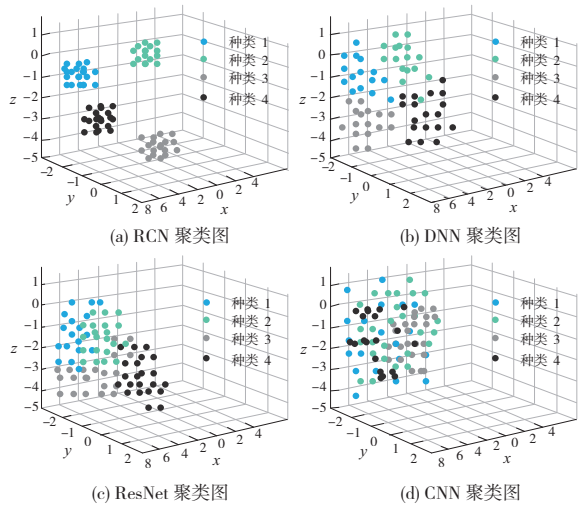


图 3　不同模型的收敛曲线图

图 3 中,RCN 在经历 22 次迭代后收敛,测试精度为 97.2%,与其他 3 种测试方法相比,残差胶囊网络模型具有更快的收敛速度,并且达到最佳的测试准确率。为了验证 RCN 模型的故障分类能力,将 RCN 模型、DNN 模型、ResNet 模型、CNN 模型对 4 种故障类型进行聚类,测试结果如图 4 所示。

从图 4 中可以看出,RCN 模型对故障的分类效果最好,与其他 3 种模型相比,RCN 模型中引入了残差块,因此具有更好的识别故障和分类效果。



(a) RCN 聚类图　(b) DNN 聚类图

(c) ResNet 聚类图　(d) CNN 聚类图

图 4　4 种模型的聚类可视化结果图

2.2　注意力胶囊网络模型旋转机械故障诊断及分析

为了测试注意力胶囊网络在旋转机械故障诊断中的性能,采用 python3.7 编程,torch5.1 深度学习框架,并采用 Adam 优化算法,学习率设置为 0.001。在 DNN 网络中设置 6 层深度神经网络,CNN 中引入一层一维卷积层,Transformer 中设置 32 个通道,特征维度为 128,CN 中引入 128 个卷积核。不同网络的收敛过程如图 5 所示。

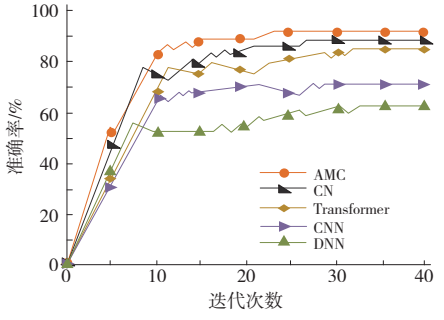


图 5　不同模型的迭代曲线图

从图 5 中可以看出:AMC 模型经过 24 次达到收敛状态,准确率为 92.5%,具有较高的测试精度和收敛速度,说明引入的 G-K 动态路由算法在信息传递中对特征的组合能力明显优于其他 4 种模型。为了验证注意力胶囊网络在实际案例中的诊断准确性,选择 WTDS 试验台提取的齿轮故障信息、滚动轴承故障信息作为数据集,不同模型的测试结果如表 1 所示。

从表 1 中可以看出,AMC 在齿轮故障中的测试准确率达到 99.2%,在轴承故障的测试中准确率达到 98.9%。这两种故障中的测试准确率明显高于其他算法,并且 AMC 在故障诊断中的稳定性为 98.8%,测试时间为 12 s,比其他方法的稳定性高,时间消耗少。因此,AMC 结合了胶囊网络和注意力机制的优点,引入的 G-K 动态路由算法能够更好地保留细节信息,提高了模型的表达能力和泛化能力。

表 1　不同模型对齿轮故障和滚动轴承故障的测试准确率表

实验方法	齿轮故障测试准确率/%	轴承故障测试准确率/%	稳定性/%	时间/s
AMC	99.2	98.9	98.8	12
CN	98.5	97.6	97.5	15
Transfomer	94.8	93.7	89.5	27
CNN	93.4	90.8	86.5	30
DNN	73.5	69.4	69.1	45

3　结语

为了实现对旋转机械中的故障诊断需求,分别在胶囊网络的基础上引入残差块和注意力机制,构建故障诊断模型。实验表明:残差胶囊网络故障诊断模型在 22 次迭代后收敛,测试精度为 97.2%;注意力胶囊网络故障诊断模型在 24 次迭代后达到收敛,准确率为 92.5%。将模型对齿轮和滚动轴承故障信息进行测试,准确率分别达到 99.2%和 98.9%。两种模型都能实现对故障精准分类,具有较强的泛化能力。此次研究尚存在不足之处,对齿轮和轴承的故障只研究了一种类型,在接下来的研究中,应考虑到多种类型的故障形式。

参考文献:

[1] 瞿红春,朱伟华,高鹏宇,等. 基于注意力循环胶囊网络的滚动轴承故障诊断[J]. 振动·测试与诊断,2022,42(6):1108-1114,1242-1243.

[2] 张娟飞. 基于 VMD 的数控机床旋转机械运行路径偏离故障检测方法[J]. 机械与电子,2022,40(11):30-34.

[3] 汪祖民,张志豪,秦静,等. 基于卷积神经网络的机械故障诊断技术综述[J]. 计算机应用,2022,42(4):1036-1043.

[4] PATURI U M R,CHERUKU S,REDDY N S. The role of artificial neural networks in prediction of mechanical and tribological properties of composites: a comprehensive review[J]. Archives of Computational Methods in Engineering,2022,29(5):3109-3149.

[5] 苏欣,高晓旭,赖复尧,等. 基于机器学习的零件加工质量预测及优化[J]. 机械制造与自动化,2022,51(5):136-139.

[6] FANG P J,CAI Y F,CHEN L,et al. A high-performance neural network vehicle dynamics model for trajectory tracking control[J]. Proceedings of the Institution of Mechanical Engineers,Part D: Journal of Automobile Engineering,2023,237(7):1695-1709.

收稿日期:2023-06-26