

# 基于注意力循环胶囊网络的滚动轴承故障诊断\*

瞿红春, 朱伟华, 高鹏宇, 王超, 周大鹏, 丁凯

(中国民航大学航空工程学院 天津, 300300)

**摘要** 针对滚动轴承工作工况复杂、载荷大及测得的振动信号信噪比(signal-to-noise ratio, 简称 SNR)低的特点, 提出了一种利用注意力循环机制(attention recurrent, 简称 AR)构建数字胶囊并与胶囊网络(capsule network, 简称 Caps)相融合的微弱故障诊断模型。首先, 在构建初级胶囊时引入双向长短期记忆网络(bidirectional long short time memory neural network, 简称 Bi-LSTM), 对时频图中的时序特征进行提取, 并建立胶囊间的非线性关联; 其次, 引入注意力循环机制构建数字胶囊, 提高时频图中不同时间和频带的能量强度变化的影响力; 然后, 通过 3D 卷积与动态路由机制构建的数字胶囊进行自适应融合, 实现特征的多样提取; 最后, 利用 softmax 分类器将融合特征映射到输出层, 实现高噪声环境下的滚动轴承故障诊断。结果表明, 该方法对小样本、低信噪比的微弱故障信号较其他诊断模型有更高的诊断精度, 并能够有效减小过拟合问题。使用不同负载下的数据做测试集验证了该模型有较强的泛化能力。

**关键词** 智能故障诊断; 胶囊网络; 注意力机制; 滚动轴承

**中图分类号** TH17; TH113.1

## 引言

滚动轴承是旋转机械转子的关键支撑部件, 长期处于复杂、恶劣的工作环境中, 具有高故障率、易损坏的特点。由于轴承所在位置特殊, 无法对轴承状况进行直接观测和诊断, 所以目前普遍采用对振动信号进行分析的方法实现滚动轴承的故障诊断。近年来, 得益于机器学习的强大拟合能力, 许多学者将机器学习应用于故障诊断领域中。徐冠基等<sup>[1]</sup>对时间序列进行相空间重构, 结合采用简化粒子群算法优化改进的双子支持向量机进行轴承故障诊断。文献[2-3]对轴承振动信号进行特征提取并利用反向传播神经网络(back propagation neural network, 简称 BPNN)和径向基网络进行分析和检测。虽然这些研究都取得了较好的结果, 但是浅层神经网络容易陷入局部最优解, 收敛速度较慢, 滚动轴承往往是在交替工况下工作, 测得的故障信号特征微弱并且易被调制呈现出典型的非线性与非平稳性, 难以通过浅层特征进行完整描述。随着深度学习技术在图片识别、计算机视觉及自然语言处理方面的成功运用, 深度学习技术也被引入滚动轴承故障诊断领域。曲建岭等<sup>[4]</sup>将原始时域信号与卷积神经网络

(convolution neural network, 简称 CNN)相结合, 实现了滚动轴承的自适应诊断。文献[5-9]将振动信号的时频分析与卷积神经网络相结合, 利用 CNN 能够自动学习数据内部特征的特性, 将时频分析提取出的特征映射到目标空间, 实现故障的诊断。文献[10]将降噪自编码网络(denoising autoencoder, 简称 DAE)与深度置信网络(deep belief networks, 简称 DBNs)相融合, 分别利用 DAE 学习浅层特征和 DBNs 学习深层特征的优势自动对振动信号进行自适应特征提取, 实现了滚动轴承故障智能诊断。虽然以上深度学习模型在完成轴承故障诊断任务时都有较好的表现, 但仍存在不足。文献[5-9]中 CNN 的池化层都是采用最大池化, 这样虽然可以增强网络对于平移和旋转的时不变性, 增强网络的泛化能力, 但却丢失了特征图层中的其他特征<sup>[11]</sup>。传统卷积网络将提取好的深层特征通过全连接层映射到目标空间, 但是由于全连接层的参数过多, 容易发生拟合现象<sup>[12]</sup>。文献[10]使用的 DAE 和 DBNs 的融合网络虽然能够实现故障特征的自动提取, 但是 DAE 和 DBNs 网络本身存在计算量大、诊断速度慢的缺点。

为了克服 CNN 导致特征丢失、全连接层引发过

\* 中国民航大学科研基金资助项目(05yk08m); 中央高校基本科研业务费资助项目(ZXH2010D019)

收稿日期: 2020-08-06; 修回日期: 2020-11-13

拟合现象等缺陷, Sabour 等<sup>[13]</sup>提出了胶囊网络(CapsNet)。胶囊网络使用向量进行网络运算,每个向量都表征1个胶囊,向量的每个维度都代表着某个实体的一种特征,所以胶囊网络模型的特征表达更丰富。Zhu 等<sup>[11]</sup>利用凯斯西储大学及帕德博恩大学的轴承数据库数据对所提出的改进胶囊网络模型进行验证,验证精度达到了97.15%。Chen 等<sup>[14]</sup>验证了胶囊网络在不同载荷、不同背景噪声下对轴承故障诊断有着良好的诊断效果以及较强的鲁棒性。袁路生等<sup>[15]</sup>利用凯斯西储大学的轴承数据库数据对胶囊网络模型的诊断能力进行验证,精度达到了99.93%。Wang 等<sup>[16]</sup>将胶囊网络应用于风力机齿轮箱的故障识别,模型的平均分类精度达到94%。Huang 等<sup>[17]</sup>在五速汽车变速器上采集了由轴承故障(内圈缺陷)、齿轮故障(缺齿缺陷)和复合故障(轴承和齿轮故障)组成的齿轮箱数据集,验证了胶囊网络在旋转机械的故障诊断中的良好表现。

笔者提出一种注意力循环机制与胶囊网络相结合的滚动轴承微弱故障诊断模型。首先,通过嵌入Bi-LSTM来重构主胶囊层,使得每个主胶囊之间都建立非线性关系,提高网络的分类性能;其次,在主胶囊层和数字胶囊层之间并行嵌入注意力循环机制,利用路由机制和注意力循环机制2种不同的方式构建数字胶囊,增加特征提取的多样性;最后,通过三维卷积运算对路由机制的数字胶囊和注意力机制的数字胶囊实现自适应融合并映射到目标空间,最终实现基于Caps-AR和连续小波变换(continuous wavelet transform,简称CWT)的故障诊断方法,并通过对不同故障的滚动轴承进行故障诊断实验,证明了所提模型的可行性与有效性。

## 1 基本理论

### 1.1 胶囊网络

胶囊网络主要由两部分组成,即初级胶囊层和数字胶囊层。其中,初级胶囊层包括卷积层及初级胶囊构建层。卷积层的计算公式为

$$a_j^l = f\left(\sum_{i \in h_j} a_i^{l-1} * w_{ij}^l + b_j^l\right) \quad (1)$$

其中: $a_j^l$ 为第 $l$ 层第 $j$ 个元素; $h_j$ 为 $l-1$ 层的第 $i$ 个卷积面; $a_i^{l-1}$ 为卷积面中的数值; $w_{ij}^l$ 为对应卷积面的卷积核; $b_j^l$ 为阈值;\*表示卷积; $f(\cdot)$ 为激活函数。

$f(\cdot)$ 通常选择线性修正单元(rectified linear unit,简称ReLU),其表达式为

$$f(x) = \max(0, x) \quad (2)$$

胶囊网络在构建初级胶囊时将卷积层提取出的标量故障特征排列为向量故障特征,生成初级胶囊 $u_i$ 。

有别于CNN网络的池化层,数字胶囊层与初级胶囊层之间的传递是通过动态路由机制实现的。动态路由机制的传递过程如下。

1) 利用初级胶囊对故障特征的类别进行预测,生成实例预测胶囊,转换公式为

$$\hat{u}_{j|i} = W_{ij} u_i \quad (3)$$

其中: $i$ 为初级胶囊标号; $j$ 为数字胶囊标号; $u_i$ 为初级胶囊层输出的第 $i$ 个初级胶囊; $\hat{u}_{j|i}$ 为实例预测胶囊; $W_{ij}$ 为权值矩阵。

2) 计算实例预测胶囊与数字胶囊之间的耦合系数,并对所有实例预测胶囊进行加权求和得到数字胶囊,计算式为

$$c_{ij} = \exp(b_{ij}) / \sum_j \exp(b_{ij}) \quad (4)$$

$$s_j = \sum_i c_{ij} \hat{u}_{j|i} \quad (5)$$

其中: $b_{ij}$ 为实例预测胶囊 $\hat{u}_{j|i}$ 与数字胶囊 $s_j$ 之间的先验连接权重, $b_{ij}$ 的初始值都设为0; $c_{ij}$ 为耦合系数,通过式(5)得到,且 $\sum c_{ij} = 1$ 。

3) 胶囊网络利用数字胶囊的模长作为分类概率,先利用式(6)的squash函数将数字胶囊 $s_j$ 的模长压缩至 $[0, 1]$ ,得到数字胶囊层最终输出数字胶囊 $v_j$ ;再利用式(7)对先验连接权重 $b_{ij}$ 进行更新

$$v_j = \text{squash}(s_j) = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \quad (6)$$

$$b_{ij} = b_{ij} + \hat{u}_{j|i} \cdot v_j \quad (7)$$

### 1.2 双向长短时记忆网络

在一般的胶囊网络中,初级胶囊之间都是相互独立的。滚动轴承振动信号的小波时频图中包含了时序信息,所以在相邻区域所提取的特征都可以用来对当前区域进行评定。标准的LSTM网络只考虑了顺序的输入,所以笔者采用Bi-LSTM网络对卷积后所得的初级胶囊进行重构,以建立初级胶囊之间的非线性关联。

Bi-LSTM网络结构包含正序和反序2个独立的隐含层,每个隐含层都由若干个LSTM神经元构成,Bi-LSTM结构如图1所示。2个隐含层具有不同的权值阈值参数,其中: $C$ 为LSTM神经元,用于记录长期信息; $a$ 为隐含层输出,用于记录短期信息,且 $a$ 会传递到相邻LSTM神经元。

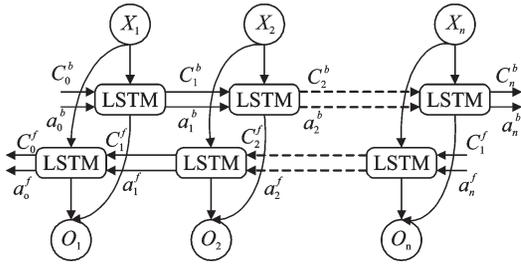


图1 Bi-LSTM结构图

Fig.1 Structure diagram of Bi-LSTM

## 2 网络结构设计

为了在强噪声下实现滚动轴承的微弱故障诊断,笔者提出注意力循环机制和胶囊网络相融合的滚动轴承故障诊断模型。所提的AR-Caps含噪信号故障诊断模型如图2所示,共分为4个部分:信号预处理、初级胶囊构建层、数字胶囊构建层和softmax分类器。信号预处理阶段包括加噪、生成小波时频图以及切分训练集和测试集。

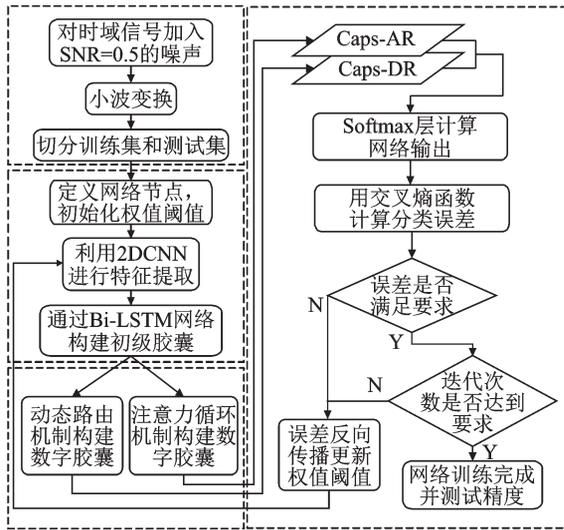


图2 AR-Caps含噪信号故障诊断模型

Fig.2 AR-Caps fault diagnosis model with noisy signals

### 2.1 初级胶囊构建层

构建初级胶囊时采用2层二维卷积网络对压缩后的灰度时频图进行特征提取,第1层卷积层采用64个尺度为 $9 \times 9$ 、步长为1的卷积核进行特征提取,第2层卷积核采用128个尺度为 $9 \times 9$ 、步长为2的卷积核进行特征提取,2个卷积层都采用ReLU激活函数。利用提取出的标量故障特征数据构建出1024个 $8 \times 1$ 的初级胶囊向量锥形,再通过输出为100维的Bi-LSTM网络对初级胶囊锥形进行重构,得到1024个 $100 \times 1$ 的最终初级胶囊向量。

### 2.2 数字胶囊构建层

#### 2.2.1 注意力循环机制

本研究的数字胶囊采用2种方式构建:①采用传统的动态路由机制构建数字胶囊,最终得到4个25维的用于故障诊断的数字胶囊;②采用注意力循环机制构建数字胶囊。这种并行模块采用不同方式进行特征提取,在提高特征多样性的同时也避免了传统神经网络在增加深度时产生的梯度消失的问题。动态路由机制能够保留所有的特征,并通过耦合系数准确反映整体与局部的关系,但是却无法有效抑制无用特征。注意力机制可以从大量信息中筛选出对当前任务更关键的信息,并且抑制无用的信息<sup>[18]</sup>。因此,笔者将路由机制的先验耦合系数与注意力机制的后验耦合系数相融合,提出注意力循环机制。通过先验耦合系数建立时频图中高能区之间的关系以提高实例预测胶囊与数字胶囊之间相似度,同时利用注意力机制提高时频图中不同时间和频带的能量强度变化的影响力,抑制不相关特征对模型的干扰,提高诊断效果。注意力循环机制的迭代过程如下,其中:输入层数为1,初级胶囊,迭代次数为 $r$ ;输出层数为 $l+1$ ,数字胶囊为 $v_i$ 。

$$1) \text{ Initialization: } b_{ij} = 0$$

$$2) \text{ For } k \text{ in range}(r)$$

$$3) c_{ij} = \text{softmax}(b_{ij})$$

$$4) \hat{u}_{j|i} = W_{ij} u_i$$

$$5) h_{j|i} = c_{ij} \odot \hat{u}_{j|i}$$

$$6) H_{j|i} = \tanh(W_{sj} h_{j|i} + b_{sj})$$

$$7) \alpha_i = \frac{\exp(H_{j|i})}{\sum_i \exp(H_{j|i})}, \sum_i \alpha_i = 1$$

$$8) s_j = \sum_i \alpha_i H_{j|i}$$

$$9) v_j = \text{squash}(s_j)$$

$$10) b_{ij} = b_{ij} + \hat{u}_{j|i} v_j$$

上述过程中:步骤3~5用来计算先验耦合系数,利用4个维度为 $100 \times 25$ 的权值矩阵 $W_{ij}$ 生成实例预测胶囊 $\hat{u}_{j|i}$ ,维度为 $1024 \times 25$ ,再利用先验耦合系数计算数字胶囊层的输入胶囊 $H_{ij}$ ;步骤6~8利用注意力机制的隐层变量及激活函数计算后验耦合系数,完成实例预测胶囊所提取特征的加权融合,所输出的初步数字胶囊 $s_j$ 的维度为 $1 \times 25$ ;步骤9对 $s_j$ 的模长进行压缩得到数字胶囊 $v_j$ ;步骤10对先验耦合系数进行更新,最终得到4个25维的用于故障诊断的数字胶囊。

2.2.2 自适应特征融合

2个独立数字胶囊的拼接方式为图2中的Caps-AR和Caps-DR,最终得到维度为 $4 \times 25 \times 2$ 的拼接特征。采用3D卷积实现2个独立数字胶囊的自适应融合,并通过softmax分类器得到维度为 $4 \times 1$ 的网络输出。该方法能够在网络迭代的过程中自适应地改变网络提取特征在网络输出的比重。路由端所得特征和注意力循环端所得特征进行融合及预测的过程为

$$\text{Caps}^l = \sum_j \sum_i a_{ij}^l \omega_{ij} \quad (8)$$

$$y = \text{softmax}(\text{Caps}) \quad (9)$$

其中: $\text{Caps}^l$ 为第 $l$ 个拼接胶囊; $a_{ij}^l$ 为第 $l$ 个拼接胶囊的第 $i$ 行、第 $j$ 列的特征量; $\omega_{ij}$ 为特征融合时的权值。

式(8)、式(9)表示将动态路由端和注意力循环端所提取出的特征进行自适应的加权融合。

3 AR-Caps模型故障诊断

3.1 数据来源

本研究使用的振动数据来自美国凯斯西储大学的开放数据库<sup>[19]</sup>,轴承实验装置如图3所示。

该实验以电机驱动端的轴承作为诊断对象,分别在测试轴承的内圈、滚动体以及外圈上采用电火花的加工方式引入单点缺陷来模拟轴承的3种故障,损伤尺度分别为0.177 8 mm,0.355 6 mm以及0.533 4 mm,尺度的变化表明故障由微弱逐渐增强。以12 kHz采样频率采集不同负载下的轴承振动数据,用于实验分析。

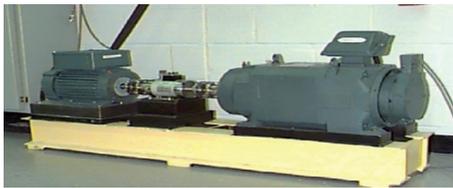


图3 轴承实验装置  
Fig.3 Bearing test device

为了实现滚动轴承的微弱故障诊断,本研究实验样本选取损伤尺度最小的0.177 8 mm、负载为735.5 W下的轴承故障样本用于模型训练。笔者将采集到的振动信号进行分割,每个样本包含512个采样点。为了模拟轴承在实际运行中的强噪声干扰,笔者采取在振动信号中加入SNR=0.5的高斯白噪声。每种轴承状态随机抽取200个样本用于模型训练,25个样本用于模型精度测试。

时域信号并不能很好地表征信号的特点,为了

能够挖掘更深层的特征,笔者采用尺度为256的cmor3-3小波基对低信噪比信号进行连续小波变换,生成小波时频图,如图4所示。由图可知,加噪信号的小波时频图高能区由原来的低频和中频区域分散到整个时频区域,使得通过小波时频图已经无法直观地进行故障诊断。将加噪后的时频图压缩成 $32 \times 32$ 大小的灰度图,最终得到不同状态下的时频图各225个,随机选取其中的800个作为训练样本,100个作为测试样本。

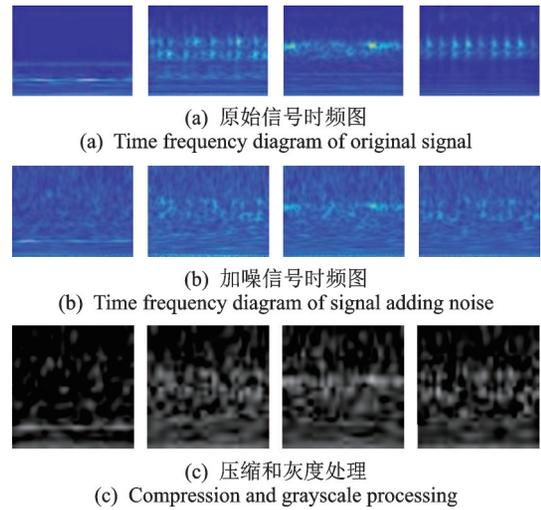


图4 小波时频图

Fig.4 Wavelet time-frequency diagram

对每种故障类型采用独热(one-hot)编码,即正常轴承表示为 $[1, 0, 0, 0]$ ,内圈故障表示为 $[0, 1, 0, 0]$ ,滚动体故障表示为 $[0, 0, 1, 0]$ ,外圈故障表示为 $[0, 0, 0, 1]$ 。

3.2 网络的训练与测试

为了保证模型诊断的精度及稳定性,降低实验的偶然性,本研究用3种不同采样方式生成的训练集与测试集,每组数据集进行2次实验,共6次实验结果的平均值对网络模型的诊断性能进行评估。6次仿真实验结果分别为98%,99%,96%,98%,95%和99%,平均正确率为97.5%。模型精度曲线如图5所示,模型损失曲线如图6所示。

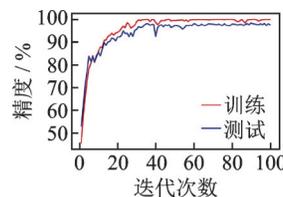


图5 模型精度曲线

Fig.5 Model accuracy curve

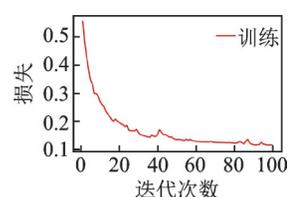


图6 模型损失曲线

Fig.6 Model error curve

结合图 5 和图 6 可以看出:AR-Caps 模型开始训练在 20 次迭代之后网络的测试精度达到 90% 以上,误差降到了 0.2 左右;在 50 次迭代之后网络的测试精度存在小幅波动,但是总体趋于平稳;最后 10 次迭代精度保持在 97.5% 左右。

### 3.3 与其他模型的对比

为评估模型精度,将所提模型与其他 6 种智能诊断模型的 6 次仿真结果进行对比,如图 7 所示。其中:AR-Caps 为笔者所提诊断模型;CWT-Caps 为连续小波变换+胶囊网络模型;CWT-CNN 为连续小波变换+CNN 模型;STFT-CNN 为短时傅里叶变换(short time Fourier transform, 简称 STFT)+CNN 模型;TDS-CNN 为低信噪比的时域信号(time domain signal, 简称 TDS)+CNN 模型;CWT-DNN 为小波时频图+深层 BP 神经网络;BPNN 为对低信噪比信号提取峰值、均值、均方根、方差、偏度和峭度<sup>[20]</sup>等指标,再结合 BPNN 的模型。

由图 7 可以看出:AR-Caps 测试精度的中位线明显高于其他模型的中位线,并高于 STFT-CNN 和 TDS-CNN 的上四分位线,与 CWT-Caps 和 CWT-CNN 的上四分位线基本持平,这表明该模型半数以上的测试精度要高于其他模型;AR-Caps 模型的中位线至上四分位线之间的距离为所有模型中最短距离之一,这说明该模型的稳定性更强,受不同采样方式的影响更低,而与其距离相同的 TDS-CNN 的整体精度低于该模型。

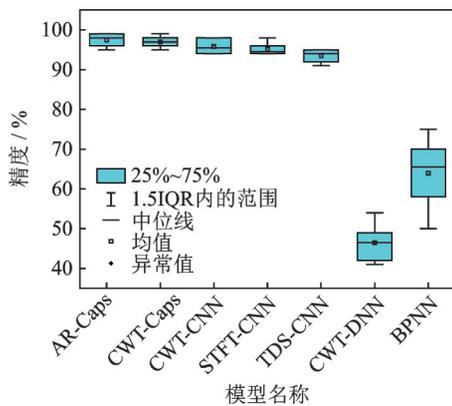


图 7 6 次仿真结果对比

Fig.7 Comparison of six simulation results

图 8 为多个模型 6 次仿真的平均结果,各模型性能对比见表 1。笔者利用测试精度与训练精度的比值来衡量模型的拟合程度。由图 8 可以看出:BPNN 模型收敛速度慢,结合表 1 中的训练精度 66.33%、测试精度 63.00% 可知,该模型受到特征数

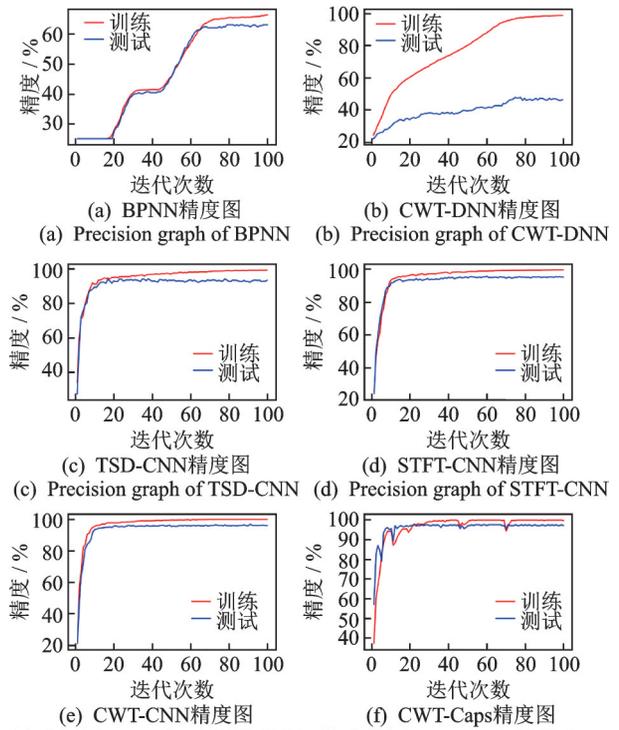


图 8 多个模型 6 次仿真的平均结果

Fig.8 Average results of six simulations of multiple models

表 1 各模型性能对比

Tab.1 Performance comparison of models %

模型	训练精度	测试精度	召回率	$F_1$	拟合程度
BPNN	66.33	63.00	63.00	57.42	94.98
CWT-DNN	98.56	46.50	46.50	45.73	47.18
TDS-CNN	99.37	93.50	93.38	90.70	94.09
STFT-CNN	99.50	95.17	95.29	95.30	95.65
CWT-CNN	99.75	95.83	95.88	95.82	96.07
CWT-Caps	99.61	97.00	97.04	96.99	97.38
AR-Caps	99.92	97.50	97.65	97.33	97.58

的限制而无法对低信噪比的振动信号进行区分;采用时频图作为输入的 DNN 模型出现了严重的过拟合,且精度表现最差;CNN 模型的收敛速度较快,并且能够很快进入稳定阶段,但其进入稳定阶段后精度无法继续提升,如 STFT-CNN 和 CWT-CNN 一直稳定在 95.17% 和 95.83% 左右,说明 CNN 模型的特征提取能力有限,从训练集中提取出的特征不足以对测试集中的数据进行识别。表 1 中 TDS-CNN 模型的测试精度为 93.50%,训练精度为 99.37%,两者的比值为 94.09%,这说明该模型出现了明显的过拟合现象。从胶囊网络模型的训练精度曲线可以看出,所有的 CapsNet 模型都会出现波谷,这说明模型从训练集中学习了新的特征,并且波谷

之后训练精度都会呈现提升趋势。虽然 CapsNet 模型前期波动较大,但最后都会达到一个相对稳定的精度值。将笔者所提模型与其他 CapsNet 模型相比,所提模型的波动更小,精度更高。

由表1可以看出:胶囊网络诊断模型的诊断精度比浅层卷积神经网络诊断模型至少高出1.17%,比BP模型至少高出34.5%;CNN诊断模型精度和召回率的调和平均数 $F_1$ 值都低于96%;BP模型的 $F_1$ 值都低于60%。这说明在小样本下胶囊网络的综合性能更好,能够提取出更多的特征,更适用于小样本低信噪比的轴承故障诊断中。本研究模型的诊断精度为97.5%,在所有胶囊网络模型中最高,其拟合程度比其他模型高出0.5%~50.4%,且召回率和 $F_1$ 值分别达到97.65%和97.33%。这说明注意力循环机制能够为时频图中含有故障信息的高能频带赋予更高的权重,可有效识别不同故障类型高能频带之间的分布区别。AR-Caps模型的测试精度与训练精度的比值为97.58%,说明其能够有效减轻过拟合现象。

### 3.4 模型适应性分析

在实际运行过程中,轴承常在不同负载下工作,为了进一步讨论所提模型在工程实际应用中的适用性与可行性,将测试集换为负载为0,1 471和2 206 W、损伤尺度为0.177 8 mm的轴承状态数据,测试样本的提取方式与训练集一致。将每种负载下的测试集样本总数增加至400个,不同负载下的仿真结果对比如图9所示。

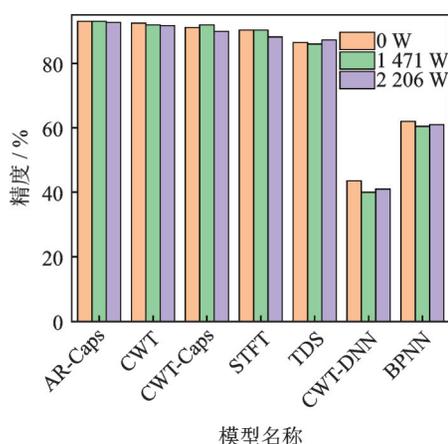


图9 不同负载下的仿真结果对比

Fig.9 Comparison of simulation results under different loads

由图9可知,当负载发生变化时,CWT-AR-Caps模型的诊断精度仍然可以达到93%,93%及92.5%,明显高于其他模型。该模型诊断精度的方差仅为 $8.33 \times 10^{-6}$ ,远低于其他模型,说明笔者所提

模型在不同负载下有更强的稳定性,能够从低信噪比振动信号的小波时频图中学习到轴承故障中的鲁棒性特征,使得模型有较强的泛化能力。

## 4 结 论

1) 针对CNN中池化层存在的特征丢失问题,提出采用动态路由机制实现低级胶囊与高级胶囊之间连接,通过赋权的形式体现出不同特征的贡献度,进而保留所有特征。

2) 针对小波时频图中包含的时域特征,利用Bi-LSTM网络对初级胶囊雏形进行重构,建立初级胶囊之间的非线性关联。

3) 针对小波时频图中频带所含能量的不同,提出注意力循环机制增强对含有故障特征的高能频带的识别能力,提高了模型对滚动轴承故障时频特征参数的提取能力。

4) 利用3D卷积运算实现数字胶囊的自适应融合,减少了特征提取层与输出层之间的连接权值,减少了模型的过拟合问题。

5) 通过仿真实验验证,相比于卷积神经网络模型及胶囊网络模型,本研究提出的模型在小样本、低信噪比的微弱故障信号上有着更好的诊断能力,对不同负载的轴承信号有着更好的泛化能力。

## 参 考 文 献

- [1] 徐冠基,曾柯,柏林.基于SPSO优化Multiple Kernel-TWSVM的滚动轴承故障诊断[J].振动、测试与诊断,2019,39(5):973-979.  
XU Guanji, ZENG Ke, BO Lin. Rolling bearing fault diagnosis based on SPSO optimization Multiple Kernel-TWSVM[J]. Journal of Vibration, Measurement & Diagnosis, 2019, 39(5):973-979. (in Chinese)
- [2] 金小强,李新民,陈焕,等.基于神经网络的直升机自动倾斜器轴承故障诊断方法[J].南京航空航天大学学报,2016,48(2):230-237.  
JIN Xiaoliang, LI Xinming, CHEN Huan, et al. Fault diagnosis method of helicopter swash-plate bearing based on neural networks[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2016, 48(2): 230-237. (in Chinese)
- [3] 李萌,陆爽,陈岱民.基于小波神经网络的滚动轴承智能故障诊断系统[J].仪器仪表学报,2005,26(8):1489-1490.  
LI Meng, LU Shuang, CHEN Daimin, et al. An intelligent fault diagnosis system of rolling bearing based on

- wavelet neural network[J]. Chinese Journal of Scientific Instrument, 2005, 26(8):1489-1490. (in Chinese)
- [4] 曲建岭, 余路, 袁涛, 等. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法[J]. 仪器仪表学报, 2018, 39(7):134-143.  
QU Jianlin, YU Lu, YUAN Tao, 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. (in Chinese)
- [5] 陈仁祥, 黄鑫, 杨黎霞, 等. 基于卷积神经网络和小波变换的滚动轴承故障诊断[J]. 振动工程学报, 2018, 31(5):161-169.  
CHEN Renxiang, HUANG Xin, YANG Lixia, et al. Rolling bearing fault identification based on convolution neural network and discrete wavelet transform[J]. Journal of Vibration Engineering, 2018, 31(5):161-169. (in Chinese)
- [6] 李恒, 张氢, 秦仙蓉, 等. 基于短时傅里叶变换和卷积神经网络的轴承故障诊断方法[J]. 振动与冲击, 2018, 37(19):132-139.  
LI Heng, ZHANG Qing, QIN Xianrong, et al. Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network [J]. Journal of Vibration and Shock, 2018, 37(19):132-139. (in Chinese)
- [7] 袁建虎, 韩涛, 唐建, 等. 基于小波时频图和CNN的滚动轴承智能故障诊断方法[J]. 机械设计与研究, 2017, 33(2):93-97.  
YUAN Jianhu., HAN Tao, TANG Jian, et al. An approach to intelligent fault diagnosis of rolling bearing using wavelet time-frequency representations and CNN [J]. Machine Design and Research, 2017, 33(2):93-97. (in Chinese)
- [8] ZENG X Q, LIAO Y X, LI W H. Gearbox fault classification using S-transform and convolutional neural network [C]//2016 Tenth International Conference on Sensing Technology (ICST). Nanjing, China: IEEE, 2016:1-5.
- [9] LIU C, CHEN G, CHEN X H, et al. Planetary gears feature extraction and fault diagnosis method based on VMD and CNN[J]. Sensors, 2018, 18(5):1523-1542.
- [10] 王应晨, 段修生. 深度学习融合模型在机械故障诊断中的应用[J]. 振动、测试与诊断, 2019, 39(6):1271-1276.  
WANG Yingchen, DUAN Xiusheng. Improved deep learning fusion model in fault diagnosis[J]. Journal of Vibration, Measurement & Diagnosis, 2019, 39(6):1271-1276. (in Chinese)
- [11] ZHU Z Y, PENG G L, CHEN Y H, et al. A Convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis[J]. Neurocomputing, 2019, 323:62-75.
- [12] ZHANG S W, ZHANG S B, ZHANG C, et al. Cucumber leaf disease identification with global pooling dilated convolutional neural network[J]. Computers and Electronics in Agriculture, 2019, 162:422-430.
- [13] SABOUR S, FROSST N, HINTON G E. Dynamic routing between capsules [EB/OL]. (2017-11-07). <https://arxiv.org/pdf/1710.09829>
- [14] CHEN T Y, WANG Z H., YANG X, et al. A deep capsule neural network with stochastic delta rule for bearing fault diagnosis on raw vibration signals[J]. Measurement, DOI: 10.1016/j. measurement. 2019. 1068570
- [15] 袁路生, 刘东东. 多级神经网络的轴承故障诊断研究[J]. 计算机工程与应用, 2020, 56(7):193-199.  
ZHONG Lusheng, LIU Dongdong. Research on bearing fault diagnosis of multi-level neural network[J]. Computer Engineering and Applications, 2020, 56(7):193-199. (in Chinese)
- [16] WANG Z J, ZHENG L K, DU W H, et al. A novel method for intelligent fault diagnosis of bearing based on capsule neural network[J]. Complexity, 2019, 2019:1-17.
- [17] HUANG R Y, LIAO Y X, ZHANG S H, et al. Deep decoupling convolutional neural network for intelligent compound fault diagnosis[J]. IEEE Access, 2019, 7:1848-1858.
- [18] 王金甲, 纪绍男, 崔琳, 等. 基于注意力胶囊网络的家庭活动识别[J]. 自动化学报, 2019, 45(11):2199-2204.  
WANG Jinjia, JI Shaonan, CUI Lin, et al. Domestic activity recognition based on attention capsule network[J]. Acta Automatica Sinica, 2019, 45(11):2199-2204.
- [19] LOPARO K A. Bearings vibration data set[D]. Cleveland: Case Western Reserve University, 2003.
- [20] 皮骏, 马圣, 贺嘉诚, 等. 基于IGA-ELM网络的滚动轴承故障诊断[J]. 航空学报, 2018, 39(9):233-244.  
PI Jun, MA Sheng, HE Jiacheng, et al. Rolling bearing fault diagnosis based on IGA-ELM network[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(9):233-244. (in Chinese)



**第一作者简介:** 瞿红春, 男, 1971年3月生, 博士、副教授。主要研究方向为基于大数据航空发动机状态监控与故障定位、民航发动机单元体和整机健康状态的评估。曾发表《基于IBSA-Elman网络的滚动轴承故障诊断研究》(《机床与液压》2020年第48卷第16期)等论文。

E-mail: hcqu2008@126.com