振动、测试与诊断
Journal of Vibration, Measurement & Diagnosis

DOI:10.16450/j.cnki.issn.1004-6801.2020.03.020

基于 LSTM-RNN 的滚动轴承故障多标签分类方法*

池永为[1]， 杨世锡[1]， 焦卫东[2]
（1. 浙江大学机械工程学院 杭州，310027） （2. 浙江师范大学工学院 金华，321004）

摘要 为了提高长短时记忆神经网络模型(long short-term memory recurrent neural network，简称 LSTM-RNN)对滚动轴承故障分类的正确率并减少训练样本量，提出一种基于多标签 LSTM-RNN 的滚动轴承故障分类方法。首先，建立滚动轴承故障信号仿真模型，分析滚动轴承故障仿真信号频谱特征及其故障分类特点；其次，结合多标签 LSTM-RNN 模型结构特点，对滚动轴承频谱特征向量进行编码，并利用仿真故障信号验证多标签 LSTM-RNN 分类方法的有效性；最后，搭建滚动轴承故障模拟试验台，采集 3 类转速不同故障类型滚动轴承故障振动信号，并采用 3 种特征提取方法得到共 9 组试验数据，基于该数据对多标签 LSTM-RNN 分类方法和单标签 LSTM-RNN 分类方法进行对比试验。试验结果表明：多标签 LSTM-RNN 分类方法相比于单标签 LSTM-RNN 分类方法，平均分类正确率从 69.07%提高到 99.21%；在保证两种分类方法正确率相近情况下，多标签 LSTM-RNN 分类方法训练所需样本量比单标签 LSTM-RNN 分类方法平均减少 69.55%。多标签 LSTM-RNN 分类方法适用于复杂振动信号分类，对于实现快速准确的旋转机械故障诊断具有应用价值。

关键词 滚动轴承；故障分类；多标签分类；循环神经网络；长短时记忆
中图分类号 TH165.3

引 言

近年来，随着人工智能技术(artificial intelligence，简称 AI)的发展，AI 技术已被应用于旋转机械故障诊断中。利用特征提取方法对检测信号进行处理得到特征序列，并利用 AI 技术对特征序列进行模式识别，从而完成旋转机械的故障诊断[1]。滚动轴承是旋转机械设备的关键零部件，容易发生点蚀、剥落或裂纹等故障。当滚动轴承发生故障时，将直接影响旋转机械系统的正常运行甚至引发安全事故[2]。滚动轴承初期故障一般表现于在滚动轴承外圈、内圈或滚子上出现点蚀现象。当滚动部件上的点蚀位置与其他零部件发生运动接触时，将引发系统内零部件复杂的周期性冲击现象。这种周期性冲击现象在频谱上体现为序列的能量变化，频谱的序列能量分布包含着滚动轴承的故障信息。经过包络分析、谱峭度分析方法[3-4]或循环平稳分析方法[5-6]解调处理后的滚动轴承振动信号特征序列故障信息更加显，充分利用该序列所包含的故障模式信息进行模式识别，对提高故障识别准确率意义重大。

循环神经网络(RNN)适合处理考虑序列前后关系的序列识别问题。针对 RNN 训练过程中的梯度消失问题提出了长短时记忆神经网络模型(LSTM-RNN)，该模型已被应用于语义分析[7]、视频分类[8]及情感分类[9]等问题。在故障诊断领域，LSTM-RNN 被应用于航空发动机的故障诊断与预测[10]。由于滚动轴承故障信号具有复杂的调制现象，故障特征序列中的故障特征频率及其边频带包含着丰富的故障模式信息。利用 LSTM-RNN 模型具有记忆功能的特点对滚动轴承故障特征序列进行模式识别，训练模型不仅可以学习故障特征频率与边频带之间的能量分布，而且可以学习各个特征频率段之间的相互关系。

多标签分类方法适用于样本具有多种类别的分类问题，每个输入样本对应多个目标标签[11]，多标签分类方法已被应用于音乐创作[12]、动作识别[13]及临床诊断等领域[14]。为了充分利用滚动轴承故障特征信息，进一步提高 LSTM-RNN 模型的故障分类准确率并减少模型的训练样本量，笔者在已有

* 国家自然科学基金资助项目((51575497；U1809219)
收稿日期：2018-07-11；修回日期：2018-10-08

的多标签分类方法基础上，结合滚动轴承故障特征分类特点，将多标签按照权重比组合成为综合标签，并利用综合标签提高 LSTM-RNN 模型分类准确性。首先，建立滚动轴承故障信号仿真模型，分析多种滚动轴承故障特征；其次，基于滚动轴承故障特征提出多标签 LSTM-RNN 分类方法，并利用滚动轴承故障仿真信号对多标签 LSTM-RNN 分类方法进行验证；最后，搭建滚动轴承故障模拟试验平台，利用多种转速滚动轴承故障信号的多种故障特征向量验证多标签 LSTM-RNN 故障分类方法的有效性。

1 LSTM-RNN 模型

1.1 RNN 模型

RNN 模型引入了定向循环的结构，这种结构使 RNN 相比于前馈神经网络(feedforward neural network，简称 FNN)更加善于解决具有长期依赖特性的序列学习问题[15]。RNN 主要应用在处理考虑历史数据影响的序列数据。因此，不同于 FNN 神经网络模型，如图 1 所示，RNN 神经网络在隐藏层内前后单元是连接的。例如，第 i 个神经元每一个隐藏层单元接受的信息来自于第 i 个输入单元和第 $i-1$ 个隐藏层单元。这样的网络结构使 RNN 能够处理现实环境中的很多应用，如语音识别、机器翻译及动作识别等。如果将滚动轴承故障特征矢量划分为一组序列，RNN 同样具有处理滚动轴承故障分类的能力。

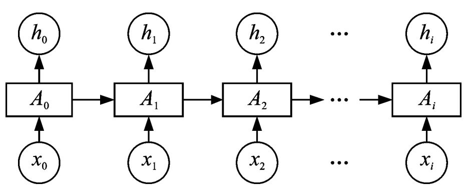


图 1 RNN 网络结构图

Fig. 1 Network structure of RNN

RNN 模型的分类原理是通过正向传播阶段和误差反向传播阶段训练 RNN 模型[16]。在正向传播阶段输入信号经过隐含层处理达到输出层，计算期望输出 H_j 和实际输出 O_j 之间的误差，均方差为

$$E = \frac{1}{2}\sum_{j}(H_j - O_j)^2 \tag{1}$$

在模型训练的反向传播阶段，输出层误差沿原路径反向传播，到达输入层。反向传播过程中修改各层神经元的连接权值，使得期望输出与实际输出的总体误差逐步达到最小。修正量为

$$\Delta W_{ji}^{k}(m+1) = -\eta\partial E/\partial W_{ij} + \alpha\Delta W_{ji}^{k}(m)$$
$$(k=1,2,\cdots,K;\ i=1,2,I) \tag{2}$$

其中：k 为输入模式的个数；i 为隐含层神经元个数；α 为动量因子；η 为学习率；m 为迭代次数。

当误差缩小到事先指定的范围或迭代次数达到设定值，则训练结束。通过以上两个过程，可以完成 RNN 模型的训练。训练结束后采用交叉熵函数进行判别期望输出序列与标签之间的相似度，相似度大的标签判别为序列所属类别。

随着 RNN 模型中序列数量的不断累加，在反馈过程中误差信息的损失量也在累加，当损失量积累到一定程度，初始信息就会退化，出现梯度消失效应。梯度消失是阻碍 RNN 应用的重要问题，当序列到达一定长度后，历史信息无法在后续的单元中进行体现，这样就丧失了 RNN 的序列学习功能。

1.2 带有窥视孔连接的 LSTM 结构

针对 RNN 的梯度消失问题，一种更加复杂的结构形式 LSTM 被引入作为 RNN 的隐藏层单元。LSTM 结构可以根据输入信息与细胞状态决定对细胞记忆覆盖、取回或保留，有效解决了 RNN 梯度消失问题[17]。为了更多地保留细胞信息，避免由于门关闭细胞信息的丢失，在 LSTM 记忆模块结构中增加窥视孔连接[18]。针对滚动轴承故障诊断的特点，需要加强序列中各个细胞状态之间的关联性，笔者采用带有窥视孔连接的 LSTM 记忆模块结构形式[19]。LSTM 记忆模块包括遗忘门、输入门和输出门结构，通过这 3 个门结构取舍输入信息并更新细胞状态。如图 2 所示，遗忘门结构由 sigmoid 层和乘法单元组成；输入门由 sigmoid 层、tanh 层和乘法单元组成；输出门由 sigmoid 层、乘法单元和 tanh 层组成。

LSTM 记忆模块内遗忘门、输入门、输出门状态可由式(3)～(8)计算[20]

$$f_t = \mathrm{sigm}(W_{x_t,f_t}x_t + W_{h_{t-1},f_t}h_{t-1} + W_{C_{t-1},f_t}C_{t-1} + b_{f_t}) \tag{3}$$

$$i_t = \mathrm{sigm}(W_{x_t,i_t}x_t + W_{h_{t-1},i_t}h_{t-1} + W_{C_{t-1},i_t}C_{t-1} + b_{i_t}) \tag{4}$$

$$g_t = \mathrm{sigm}(W_{x_t,i_t}x_t + W_{h_{t-1},i_t}h_{t-1} + W_{C_{t-1},i_t}C_{t-1} + b_{i_t}) \tag{5}$$

$$C_t = f_tC_{t-1} + i_tg_t \tag{6}$$

$$o_t = \mathrm{sigm}(W_{x_t,o_t}x_t + W_{h_t,o_t}h_{t-1} + W_{C_t,o_t}C_t + b_{o_t}) \tag{7}$$

$$h_t = o_t\tanh(C_t) \tag{8}$$

其中：x_t 为输入单元；h_t 为输出单元；f_t 为遗忘门

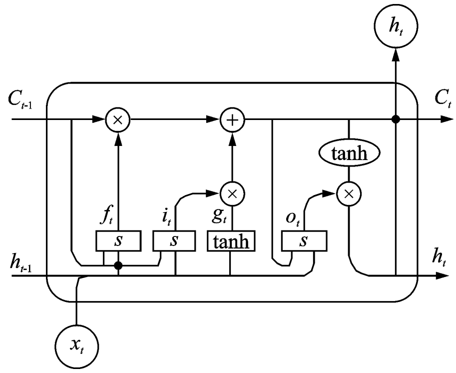


图 2　带有窥视孔连接的 LSTM 记忆模块结构图

Fig. 2　Memory module structure diagram of LSTM with peephole connection

状态;i_t 为输入门状态;o_t 为输出门状态;C_t 为细胞状态;g_t 为候选细胞状态;W 为权重系数;b 为偏置系数;sigm() 为 Sigmoid 函数;tanh() 为双曲正切函数。

带有窥视孔连接的 LSTM 记忆模块能够使前一个细胞状态参与到遗忘门、输入门和输出门的计算过程中。由式(3)～(8)所示,在 LSTM-RNN 模型不断学习过程中,输入序列中各个元素与前一个细胞状态共同作为细胞训练的输入量,使 LSTM-RNN 模型具有了记忆功能。在滚动轴承故障分类过程中,解调后的频谱特征作为主要的故障识别特征被广泛应用。滚动轴承故障信号各个特征频率间的幅值关系是判断故障类型的主要依据。因此,具有记忆功能的 LSTM-RNN 模型符合滚动轴承故障分类要求。

2　滚动轴承故障信号仿真建模与特征分析

2.1　滚动轴承故障信号仿真建模

当滚动轴承部件上的点蚀位置与其他零部件发生运动接触时,点蚀故障将引发振动信号中的周期性冲击现象。冲击频率取决于点蚀发生的零部件位置,当滚动轴承几何结构确定后,外圈、内圈、滚子等零部件上发生点蚀故障时产生的故障冲击频率可由式(9)～(11)计算[21]

$$f_{\text{out}} = \frac{N_b}{2}\left(1 - \frac{d}{D}\cos\theta\right) f_{\text{r}} \tag{9}$$

$$f_{\text{in}} = \frac{N_b}{2}\left(1 + \frac{d}{D}\cos\theta\right) f_{\text{r}} \tag{10}$$

$$f_{\text{ball}} = \frac{D}{2d}\left[1 - \left(\frac{d\cos\theta}{D}\right)^2\right] f_{\text{r}} \tag{11}$$

其中:f_{out} 为外圈故障频率;f_{in} 为内圈故障频率;f_{ball} 为滚子故障频率;N_{b} 为滚子数量;d 为滚子直径;D 为轴承节圆直径;f_{r} 为转动频率;θ 为压力角。

针对不同滚动轴承故障类型,故障信号仿真模型主要考虑由点蚀引起的冲击振荡信号和环境噪声信号,其中冲击幅值受到不同旋转体转动调制。滚动轴承故障信号仿真模型为

$$x(t) = \sum_i A_i s(t - iT - \tau_i) + n(t) \tag{12}$$

其中:$x(t)$ 为滚动轴承故障振动信号;A_i 为第 i 次冲击的幅值;$s(t)$ 为点蚀故障产生的冲击振荡;T 为 2 次冲击之间的时间间隔;τ_i 为第 i 次故障冲击时轴承滚子打滑引起的周期 T 的波动;$n(t)$ 为零均值平稳随机噪声;t 为时间。

第 i 次冲击的幅值 A_i 为

$$A_i = \cos(2\pi f_{\text{A}} t + \varphi_{\text{A}}) + C_{\text{A}} + \text{rand}n(t) \tag{13}$$

其中:f_{A} 为调制频率,为零均值随机信号;φ_{A} 和 C_{A} 为任意常数。

根据点蚀故障位置不同,f_{A} 的取值不同,外圈发生故障时 $f_{\text{A}}=0$,内圈发生故障时 $f_{\text{A}}=f_{\text{r}}$,滚子发生故障时 $f_{\text{A}}=f_{\text{c}}$,$f_{\text{c}}$ 为保持架转动频率

$$f_{\text{c}} = \frac{1}{2}\left(1 - \frac{d}{D}\cos\theta\right) f_{\text{r}} \tag{14}$$

冲击振荡 $s(t)$ 是以系统固有频率为振荡频率的振荡衰减信号

$$s(t) = \text{e}^{-Bt}\sin 2\pi f_n t \tag{15}$$

其中:f_n 为系统固有频率;B 为常数。

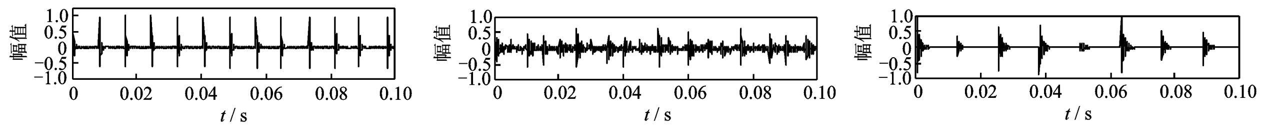
表 1 为滚动轴承故障信号仿真建模所采用的参数,利用 MB ER-12K 滚动轴承尺寸计算滚动轴承特征频率,其中 MB ER-12K 滚动轴承为试验系统中所装配的轴承型号。

表 1　故障信号仿真建模参数表

Tab. 1　Parameters table for fault signal simulation

参数	取值	参数	取值
D/mm	47.00	d/mm	7.94
N_{b}	8	f_{r}/Hz	40
θ/(°)	0	f_n/Hz	3 000
B	2 500		

利用式(12)～(15)和表 1 中参数得到一组转频为 40Hz 的滚动轴承故障仿真信号,经过归一化后的时域信号如图 3 所示,按照式(9)～(11)计算得到 3 种故障信号冲击频率。内圈故障和滚子故障仿真时域信号冲击幅值调制现象比外圈故障信号更加明显,周期性冲击及其调制现象都会在频谱上体现。



(a) 滚动轴承外圈点蚀故障仿真信号
(a) Simulated fault signal of rolling bearing with outer-race defect

(b) 滚动轴承内圈点蚀故障仿真信号
(b) Simulated fault signal of rolling bearing with inner-race defect

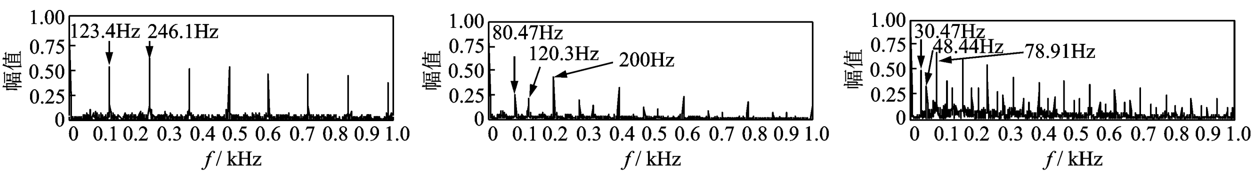
(c) 滚动轴承滚子点蚀故障仿真信号
(c) Simulated fault signal of rolling bearing with ball defect

图3 滚动轴承故障仿真信号

Fig. 3 Simulated fault signal of rolling bearings

2.2 滚动轴承故障特征分析

当滚动部件上的点蚀位置与其他零部件发生运动接触时，点蚀故障将引发周期性冲击现象。滚动轴承振动信号经频域分析后，特定故障频率的振动能量成分能够反映滚动轴承故障类型。但由于故障特征淹没在高频噪声中，滚动轴承故障信号需要经过解调处理提取故障特征，常用的解调方法有包络分析、谱峭度分析方法及循环平稳分析方法。将图3中滚动轴承故障仿真信号进行平方包络后再进行傅里叶变换，分别得到滚动轴承外圈故障、内圈故障和滚子故障信号频谱图，见图4。为了对比说明滚动轴承故障特征频率，利用式(9)～(11)和式(14)计算转频为40Hz时的故障特征频率，如表2所示。



(a) 滚动轴承外圈故障仿真信号频谱
(a) Frequency spectrum of simulatedfault signal of rolling bearing with outer-race defect

(b) 滚动轴承内圈点蚀故障仿真信号
(b) Frequency spectrum of simulatedfault signal of rolling bearing with inner-race defect

(c) 滚动轴承滚子点蚀故障仿真信号
(c) Frequency spectrum of simulatedfault signal of rolling bearing with balldefect

图4 滚动轴承故障仿真信号平方包络频谱图

Fig. 4 Frequency spectrum of square envelope signal of faulty rolling bearing

表2 滚动轴承故障特征频率表

Tab. 2 Fault characteristic frequency table of rolling bearing

Hz

特征频率	数值	特征频率	数值
转频	40	外圈故障频率	121.92
2倍转频	80	内圈故障频率	198.00
3倍转频	120	滚子故障频率	79.68
4倍转频	160	保持架频率	15.12

如图4(a)所示的外圈故障仿真信号频谱中，在外圈故障特征频率123.4Hz和其2倍频处有明显峰值；图4(b)所示的内圈故障仿真信号频谱中，在转动频率2倍频80.47Hz、3倍频120.3Hz和外圈故障频率200Hz处有明显峰值；图4(c)所示的滚子故障仿真信号频谱中，在保持架转频2倍频30.47Hz、3倍频48.44Hz和滚子故障频率78.91Hz处有明显峰值。在故障信号建模过程中考虑了轴承滚子随机打滑因素，导致特征频率在计算值附近产生一定波动。利用谱峭度滤波方法和循环平稳分析方法均可以得到故障特征向量。在试验研究中将对不同特征提取方法得到的特征向量进行机器学习并进行分类，以分类结果分析讨论特征提取方法的特点。

3 多标签LSTM-RNN滚动轴承故障分类方法

3.1 滚动轴承故障特征编码

在对滚动轴承进行故障诊断过程中，维护人员以频谱作为主要的判断依据。由式(9)～(11)得到滚动轴承故障特征频率，利用故障特征频率、转频及其倍频和边频带的幅值大小关系可以说明滚动轴承故障类型。

基于以上对滚动轴承故障诊断经验的分析，笔者提出以转频、2～6倍转频、外圈故障特征频率、内圈故障特征频率及滚子故障特征频率为中心频率、20Hz为带宽生成10个频谱序列。以此序列为LSTM-RNN模型的学习目标。这种特征编码方式既保留了特征频率点的边频带信息，又考虑了多个特征频率段之间的关系，且特征频率序列包含了滚动轴承打滑所引起特征频率随机波动信息。特征频率序列包含了故障分类所需要的信息，为LSTM-

RNN 学习模型提供充足信息保证分类成功率。

分类标签采用 L_1，L_2，L_3 及 L_4 分别作为健康状态、外圈故障、内圈故障及滚子故障。其中

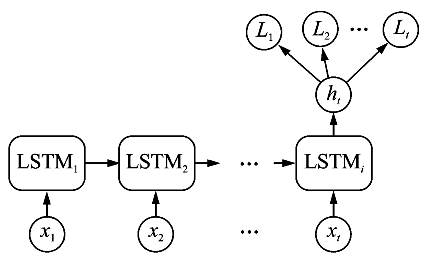
$$\begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{16}$$

多标签 LSTM-RNN 学习模型采用的标签为长度与输入序列相同的一组标签序列 $[L_{1\sim4}^{i}]$，其中 i 为序列长度。

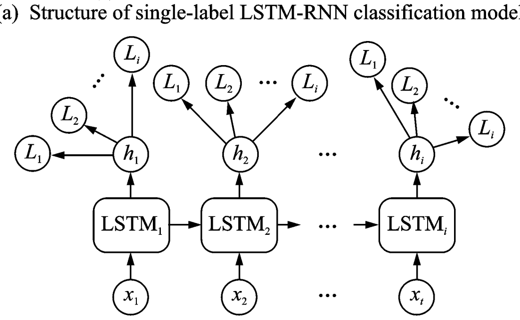
3.2　多标签 LSTM-RNN 分类模型

笔者基于滚动轴承故障特征提出一种多标签 LSTM-RNN 故障分类模型。LSTM-RNN 分类模型利用式(3)～(9)进行正向传播计算，并利用误差反向传播算法对序列中各个细胞状态进行更新，如图 5 所示。

多标签 LSTM-RNN 分类模型与单标签分类模型的区别如下：多标签模型中输入序列中每个元素都对应 1 个输出元素，每个输出元素均采用绝对值损失函数进行误差反向传播更新序列内细胞状态；而单标签模型中只在最后一个元素位置采用绝对值损失函数进行误差反向传播更新全序列细胞状态。



(a) 单标签LSTM-RNN分类模型结构
(a) Structure of single-label LSTM-RNN classification model



(b) 多标签LSTM-RNN分类模型结构
(b) Structure of multi-label LSTM-RNN classification model

图 5　LSTM-RNN 分类模型结构图
Fig. 5　Structure diagram of LSTM-RNN classification model

在 LSTM-RNN 模型训练结束后，可以利用该模型进行故障分类。将测试数据集带入模型得到输出数据，并将输出数据带入多标签分类器中进行分类。多标签分类器的操作流程是先利用 Softmax 函数对每个元素内的输出值进行归一化处理，再利用交叉熵函数判定输出值与标签之间的相似性。单个标签和输出元素之间损失值计算如下

$$\mathrm{loss}(h_i, L_i) = -\frac{1}{M}\sum_{m=1}^{M} L_i^m \log(h_i^m) + (1-L_i^m)\log(1-h_i^m) \tag{17}$$

其中：h_i 为序列中第 i 个输出值；L_i 为标签序列中第 i 个标签值；M 为每个元素中的数据量；L_i^m 为标签序列中第 i 个元素中的第 m 个数值；h_i^m 为输出序列中第 i 个元素中的第 m 个数值。

基于每个标签的损失值可以计算序列的损失值，判断序列相似性

$$\mathrm{loss}_{\mathrm{seq}} = -\frac{1}{I}\sum_{i=1}^{I} \alpha_i \mathrm{loss}(h_i, L_i) \tag{18}$$

其中：α_i 为权重系数。

α_i 由训练集数据计算得到，将训练集数据序列中的每个元素与标签序列中的每个标签进行交叉熵计算，并将交叉熵值带入 Softmax 函数，得到每个标签的权重系数 α_i 。

单标签分类方法仅采用最后一个标签进行故障分类，而多标签分类方法则通过所有标签进行故障分类，并分配了每个标签的权重比例。因此，多标签分类方法对特征向量信息利用的更加充分。

3.3　滚动轴承仿真故障信号分类验证

分别利用多标签和单标签 LSTM-RNN 分类模型对滚动轴承故障仿真信号特征向量进行分类，验证多标签 LSTM-RNN 方法的有效性。首先，将图 4中滚动轴承故障仿真信号特征向量训练集按照前述编码方法进行特征编码；其次，分别利用多标签和单标签 LSTM-RNN 分类模型对滚动轴承故障仿真信号进行分类，利用滚动轴承仿真信号测试集进行分类正确率验证。其中：训练模型参数学习率设置为 0.1；动量因子设置为 0.5；学习率变化因子设置为 0.5。

多标签 LSTM-RNN 分类模型中各标签权重比例如图 6 所示，按照此权重比例分配各个标签在序列分类中的所占比重。

为了对比多标签 LSTM-RNN 分类模型和单标签 LSTM-RNN 分类模型的分类效果，对不同样本量训练集进行分类并统计分类正确率，如图 7 所示。随着训练集样本量逐渐增加时，多标签和单标签 LSTM-RNN 分类模型的分类正确率均有所提高。利用多项式拟合图 7 中样本点，虚线型拟合曲线最

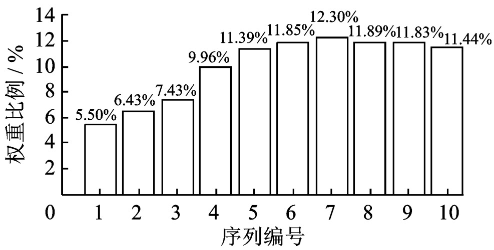


图 6 仿真故障信号各标签权重比例分配图

Fig. 6 Weight distribution of each label for the simulated fault signal

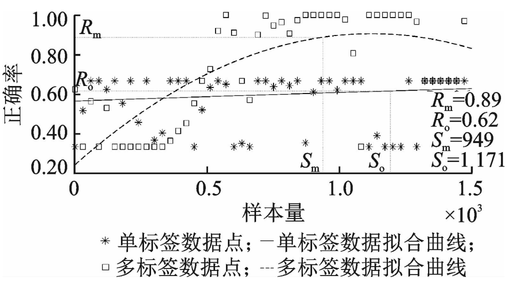


图 7 多标签和单标签 LSTM-RNN 分类正确率对比图

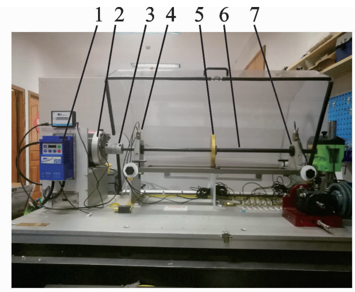
Fig. 7 Comparison of classification accuracy between multi-label and single-label LSTM-RNN

高点代表多标签 LSTM-RNN 分类模型最高正确率 $R_m=0.89$，相同方法得到单标签 LSTM-RNN 分类模型正确率最高为 $R_o=0.62$。取拟合曲线与最高正确率位置水平线交点，在交点中取最小横坐标值作为多标签和单标签分类方法达到最大正确率时所需样本量 S_m 和 S_o，其中 $S_m=949$，$S_o=1\ 171$。仿真数据集分类结果说明，多标签 LSTM-RNN 分类方法比单标签 LSTM-RNN 分类方法正确率更高，训练所需要样本量更少。

4 试验研究

4.1 试验系统搭建

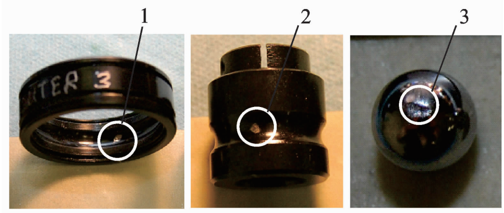
滚动轴承故障模拟试验台由电机、调速器、联轴器、转子盘、主轴和轴承组成，振动信号由加速度传感器采集，采集频率为 25.6kHz。图 8 (a)为滚动轴承故障模拟试验台的轴系结构。轴承可拆卸替换为 4 种健康状态轴承，分别为健康轴承、外圈故障轴承、内圈故障轴承和滚子故障轴承。点蚀故障发生位置分别为外圈、内圈和滚子。图 8 (b)为滚动轴承不同故障类型的故障零部件。



1-调速器; 2-电机; 3-联轴器; 4-加速度传感器; 5-转子盘; 6-主轴; 7-故障轴承

(a) 滚动轴承故障模拟试验台

(a) Test rigfor simulating rolling bearing fault



1-外圈点蚀故障; 2-内圈点蚀故障; 3-滚子点蚀故障

(b) 滚动轴承点蚀故障

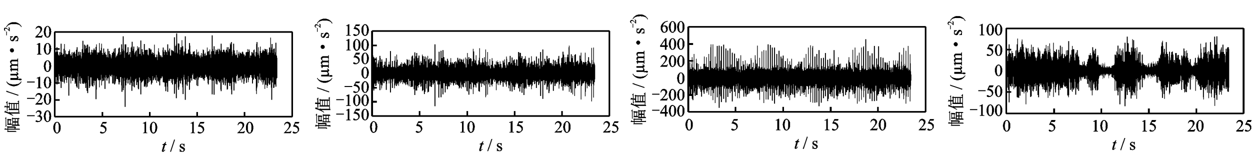
(b) Pitting failure of rolling bearings

图 8 滚动轴承故障模拟试验装置

Fig. 8 Experiment device for fault simulation of rolling bearing

分别采集健康状态轴承、外圈故障轴承、内圈故障轴承和滚子故障轴承振动信号，如图 9 所示。利用平方包络、谱峭度滤波和循环平稳方法对不同类型振动信号进行特征提取，将 3 类特征向量划分为训练集和测试集作为分类试验的数据集。

滚动轴承故障信号平方包络信号频谱如图 10 所示。健康轴承信号频谱中在转频、2 倍转频和 4 倍转频处有明显突出幅值成分；外圈故障信号频谱中在转频、2 倍转频和外圈故障频率处有明显突出幅值成分；内圈故障信号频谱中在转频、2 倍转频和内圈故障频率处有明显突出幅值成分。



(a) 健康轴承 (b) 外圈故障轴承 (c) 内圈故障轴承 (d) 滚子故障轴承

(a) Healthy bearing (b) Fault bearing with outer-race defeat (c) Fault bearing with inner-race defeat (d) Fault bearing with ball defeat

图 9 滚动轴承不同状态时域信号

Fig. 9 Time domain signals of rolling bearings in different states

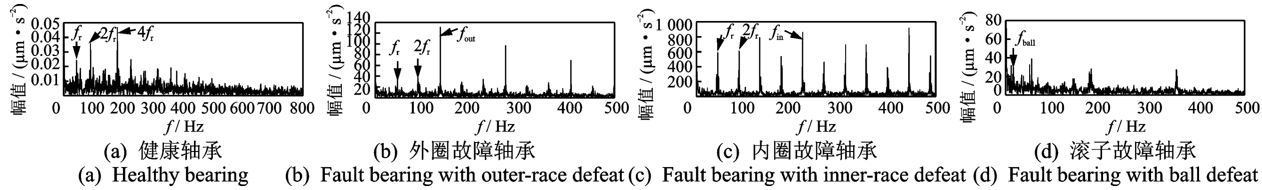


图 10　滚动轴承不同状态平方包络信号频谱

Fig. 10　Spectrum of square envelope signal of rolling bearing in different states

4.2　多标签 LSTM-RNN 分类方法试验结果

分别对实测滚动轴承故障特征向量进行编码,并利用单标签和多标签 LSTM-RNN 分类方法对特征序列进行分类试验。以转频 40Hz 的平方包络振动信号的频率特征向量数据集为例,说明分类结果的评价指标。

训练样本量为 4 000,测试样本量为 500。样本采集、特征提取、模型训练和测试等试验流程如下:

1) 分 3 次采集 90s 滚动轴承振动加速度数据,每次采集时间为 30s,采样频率为 25.6kHz;

2) 将 90s 振动加速度信号划分为 4 000 个训练样本和 500 个测试样本;

3) 对训练样本和测试样本进行数据特征提取,可得到平方包络频谱、谱峭度滤波频谱或循环平稳度;

4) 将每个样本划分为由若干特征频率段组成的一组序列;

5) 将 4 000 个样本分别代入多标签和单标签 LSTM-RNN 模型中进行训练,每隔 40 个样本进行 1 次测试,测试分类准确率;

6) 将式(18)中权重系数提取出来绘制标签权重分配图,如图 11(a)所示;

7) 将不同样本量下测试样本准确率提取出来绘制准确率变化曲线,如图 11(b)所示。

利用实测滚动轴承数据训练集分别训练多标签和单标签 LSTM-RNN 分类模型,用测试集测试不同样本量情况下的分类正确率,如图 11(a)所示。为了对比多标签和单标签 LSTM-RNN 分类方法的有效性,利用多标签分类方法最大正确率 R_m 和单标签分类方法 R_o 评估分类正确率。利用多项式拟合图 11(b)中样本点,取拟合曲线与最高正确率位置水平线交点,在交点中取最小横坐标值作为多标签和单标签分类方法达到最大正确率时所需样本量 S_m 和 S_o,并以此评估分类方法的效率。

采用不同特征频率序列进行 LSTM-RNN 模型的训练和故障分类试验。设特征序列 F_1 特征频率为 1/2 倍转频、转频、2~6 倍转频、外圈故障特征频率、内圈故障特征频率及滚子故障特征频率,共 10 段序列;设特征序列 F_2 为转频、2~5 倍转频、外圈故障特征频率、内圈故障特征频率、滚子故障特征频率、2 倍外圈故障特征频率、2 倍内圈故障特征频率及 2 倍滚子故障特征频率,共 10 段序列。特征序列 F_1 中每段序列频率范围为 20Hz,特征序列 F_2 中每段序列频率范围为 45Hz。对比 2 类不同特征序列情况,图 11 为 F_1 特征序列训练分类结果,图 12 为 F_2 特征序列训练分类结果。利用特征 F_1 和特征 F_2 序列进行多标签分类时,最大正确率相同,特征 F_2 序列在单标签分类时分类正确率会提高。本研究后续试验采用特征 F_1 序列进行试验。

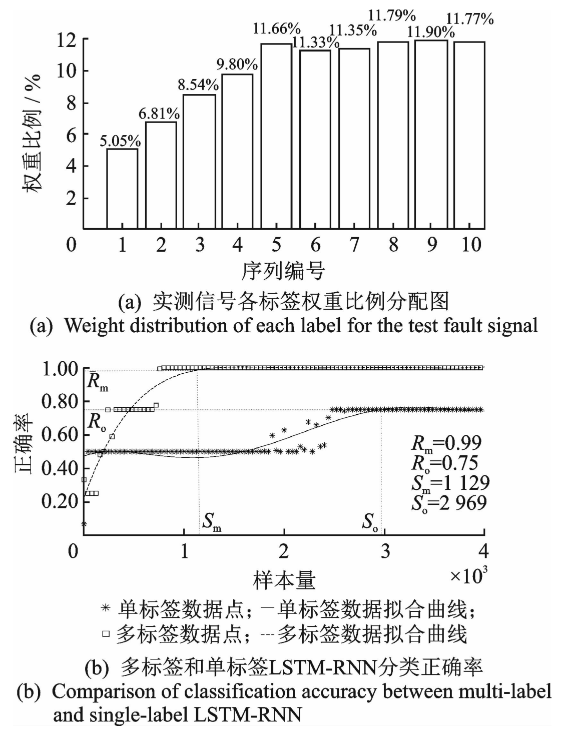
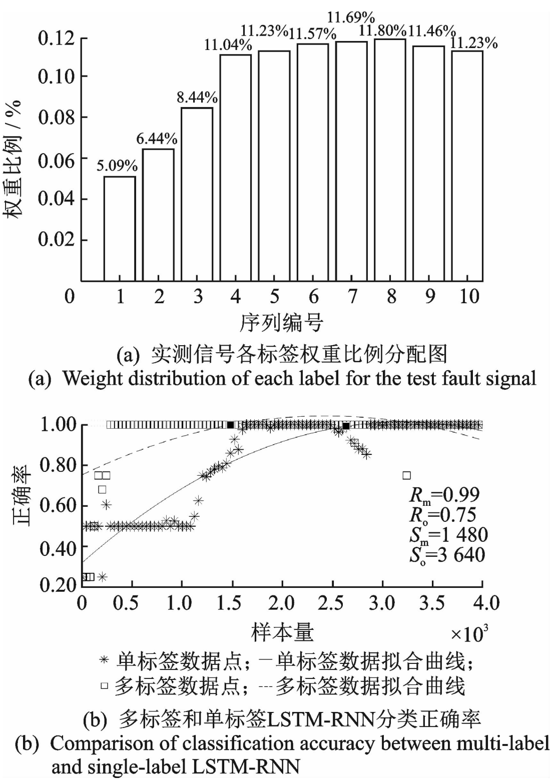


图 11　特征序列 F_1 训练模型分类结果

Fig. 11　Classification results of training model with feature sequence F_1

分别利用多标签和单标签 LSTM-RNN 分类方法对实测滚动轴承数据训练集进行故障分类试验。数据集包括 3 类转频信号,分别为 35,40 和 45Hz。



(a) 实测信号各标签权重比例分配图

(a) Weight distribution of each label for the test fault signal

(b) 多标签和单标签LSTM-RNN分类正确率

(b) Comparison of classification accuracy between multi-label and single-label LSTM-RNN

图 12 特征序列 F_2 训练模型分类结果

Fig. 12 Classification results of training model with feature sequence F_2

每种转频信号使用 3 种特征提取方法得到 3 类特征向量，分别为平方包络信号频谱、谱峭度滤波信号频谱和循环平稳度。分类试验对共计 9 组数据集分别进行训练和测试，测试结果分别以正确率和所需样本量为评价指标，见图 11。

多标签分类方法正确率 R_m 和单标签分类方法正确率 R_o 统计结果见表 3。取 9 组数据的正确率差值平均值，多标签分类方法比单标签分类方法正确率平均高 30.14%。利用多标签分类方法对由 3 种特征提取方法得到的特征向量进行分类，正确率均可接近 100%。但是如果使用单标签分类方法，基于循环平稳度的分类正确率偏低，原因在于经过循环平稳分析方法解调后得到的循环平稳度信息量相对较少。虽然特征频率相比于其他特征提取方法更加明显，但是训练样本信息量过少不利于分类模型进行学习。

多标签分类方法训练所需样本量 S_m 和单标签分类方法训练所需样本量 S_o 统计结果见表 4。取 9 组数据中正确率差 $\Delta R<20\%$ 的数据组，其多标签分类方法训练所需数据量比单标签分类方法训练所需数据量平均减少 69.55%，说明当 2 种分类方法正确率接近的情况下，多标签分类方法训练速度更快，完成训练需要样本量更少。对比 3 种特征向量所需样本量，循环平稳度数据集所需训练样本量更多。

表 3 多标签分类模型与单标签分类模型分类正确率对比统计表

Tab. 3 Accuracy of multi-label and single-label classification models %

数据类型		R_m	R_o	$\Delta R=R_m-R_o$
特征向量	转频/Hz			
平方包络信号频谱	35	99.73	81.84	17.89
	40	99.80	76.68	23.12
	45	100.00	87.20	12.80
谱峭度滤波信号频谱	35	97.34	72.58	24.76
	40	98.62	97.64	0.98
	45	98.51	79.46	19.05
循环平稳度	35	99.18	50.22	48.96
	40	99.74	50.90	48.84
	45	100.00	25.13	74.87
平均值		99.21	69.07	30.14

表 4 多标签分类模型与单标签分类模型所需样本量对比统计表

Tab. 4 Sample size of multi-label and single-label classification models

数据类型		S_m	S_o	ΔS/%
特征向量	转频/Hz			
平方包络信号频谱	35	1 245	3 997	68.85
	40	1 281	3 005	57.37
	45	761	2 509	69.67
谱峭度滤波信号频谱	35	1 553	2 161	28.14
	40	793	3 501	77.35
	45	1 129	2 997	62.33
循环平稳度	35	3 105	2 601	−19.38
	40	1 849	2 821	34.46
	45	2 417	1 665	−45.17

$\Delta S=(S_o-S_m)/S_o$。

多标签 LSTM-RNN 分类方法相比于单标签分类方法能够学习更多特征信息。滚动轴承故障分类试验结果显示其正确率更高，所需样本量更少，因此，多标签 LSTM-RNN 分类方法更加适合滚动轴承故障分类。

5 结 论

1）具有记忆功能的 LSTM-RNN 模型适合于滚动轴承故障特征序列多个元素之间的关系学习，按照每个标签的相似程度分配多个标签的决策比重，多标签分类方法更加充分利用滚动轴承特征序

列信息进行分类。

2) 在平方包络信号频谱、谱峭度滤波信号频谱和循环平稳度 3 类特征向量中,在保证相同分类正确率的前提下,循环平稳度由于数据量少导致需要更多训练样本。

3) 经过仿真数据和试验数据验证,多标签 LSTM-RNN 分类方法相比于单标签分类方法正确率平均高 30.14%。在保证 2 种分类方法正确率 20%以内的情况下,多标签 LSTM-RNN 分类方法训练所需样本量比单标签分类方法平均减少 69.55%。

参　考　文　献

[1] LIU R, YANG B, ZIO E, et al. Artificial intelligence for fault diagnosis of rotating machinery: a review[J]. Mechanical Systems & Signal Processing, 2018, 108: 33-47.

[2] HENG A, ZHANG S, TAN A C C, et al. Rotating machinery prognostics: state of the art, challenges and opportunities[J]. Mechanical Systems and Signal Processing, 2009, 23(3):724-739.

[3] WANG Y, XIANG J, MARKERT R, et al. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: a review with applications[J]. Mechanical Systems and Signal Processing, 2016, 66: 679-698.

[4] ANTONI J, RANDALL R B. The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines[J]. Mechanical Systems & Signal Processing, 2006, 20(2):308-331.

[5] SHIYANG L, NING C, PENG Y, et al. Cyclostationary approach to detect flow-induced effects on vibration signals from centrifugal pumps[J]. Mechanical Systems and Signal Processing, 2019,114:275-289.

[6] MAHVASH A, LAKIS A A. Application of cyclic spectral analysis in diagnosis of bearing faults in complex machinery[J]. Tribology Transactions, 2015, 58(6):1151-1158.

[7] PASCUAL S, BONAFONTE A. Multi-output RNN-LSTM for multiple speaker speech synthesis with α-interpolation model[C]// 9th ISCA Workshop on Speech Synthesis. Sunnyvale,CA, USA: IEEE,2016.

[8] WANG X, GAO L, SONG J, et al. Beyond frame-level CNN: saliency-aware 3-D CNN with LSTM for video action recognition[J]. IEEE Signal Processing Letters, 2017, 24(4):510-514.

[9] SUN B, WEI Q, LI L, et al. LSTM for dynamic emotion and group emotion recognition in the wild[C] // Proceedings of the 18th ACM International Conference on Multimodal Interaction. Tokyo,Japan:Association for Computing Machinery(ACM), 2016.

[10] ZHAO R, YAN R, CHEN Z, et al. Deep learning and its applications to machine health monitoring[J]. Mechanical Systems and Signal Processing, 2019, 115:213-237.

[11] DEVKAR R, SHIRAVALE S. A survey on multi-label classification for images[J]. International Journal of Computer Applications, 2017, 162(8):39-42.

[12] LIU I, RAMAKRISHNAN B. Bach in 2014: music composition with recurrent neural network[EB/OL]. [2018-06-06]. arXiv: 1412.3191, 2014.

[13] YEUNG S, RUSSAKOVSKY O, NING J, et al. Every moment counts: dense detailed labeling of actions in complex videos[J]. International Journal of Computer Vision, 2018, 126(2/4):375-389.

[14] LIPTON Z C, KALE D C, ELKAN C, et al. Learning to diagnose with LSTM recurrent neural networks [EB/OL]. [2018-06-06]. arXiv:1511.03677, 2015.

[15] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization [EB/OL]. [2018-06-06]. arXiv: 1409.2329,2014.

[16] LIU H, ZHOU J, ZHENG Y, et al. Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders[J]. Isa Transactions, 2018, 77: 167-178.

[17] MEMORY L S. Long short-term memory[J]. Neural Computation, 2010, 9(8):1735-1780.

[18] MONNER D, REGGIA J A. A generalized LSTM-like training algorithm for second-order recurrent neural networks[J]. Neural Networks, 2012, 25(1):70-83.

[19] GERS F A, SCHMIDHUBER J. Recurrent nets that time and count[C] // Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. Como,Italy:IEEE, 2000.

[20] GERS F A, SCHMIDHUBER E. LSTM recurrent networks learn simple context-free and context-sensitive languages[J]. IEEE Transactions on Neural Networks, 2001, 12(6):1333-1340.

[21] NIU L, CAO H, HE Z, et al. A systematic study of ball passing frequencies based on dynamic modeling of rolling ball bearings with localized surface defects[J]. Journal of Sound & Vibration, 2015, 357:207-232.

第一作者简介:池永为,男,1986 年 4 月生,博士生。主要研究方向为旋转机械故障诊断与健康管理。曾发表《Spectral DCS-based feature extraction method for rolling element bearing pseudo-fault in rotor-bearing system》(《Measurement》2019,Vol. 132)等论文。
E-mail:chiyw@zju.edu.cn

通信作者简介:杨世锡,男,1968 年 1 月生,博士、教授。主要研究方向为旋转机械状态监测与故障诊断、智能检测技术。
E-mail:yangsx@zju.edu.cn