

doi:10.11835/j.issn.1000-582X.2020.09.010

基于深度学习的含未知复合故障多传感器信号故障诊断

邢砾文^a, 姚文凯^a, 黄莹^b

(武警工程大学 a. 研究生大队; b. 信息工程学院, 西安 710086)

摘要:深度学习在故障诊断领域的应用已比较成熟,其中卷积神经网络(CNN, convolution neural networks)和长短时记忆网络(LSTM, long short-term memory networks)就是典型模型之一。CNN作为一种常用的多传感器信号故障诊断方法,能够获得较好的诊断效果,却无法实现未知复合故障的诊断,为了解决这个问题,提出 CNN-LSTM-FCM(fuzzy C-means)模型。LSTM对具有前后联系的时间信号更敏感,利用这个特点将 LSTM 与 CNN 相结合,实现未知信号的诊断,并通过概率分类输出实现了复合故障的解耦,CNN-LSTM-FCM 模型本身优化参数设计,进一步提高了诊断精度。使用化学过程故障测量数据进行实验,结果表明 CNN-LSTM-FCM 模型诊断准确率可达到 97.15%,优于 CNN 模型和 LSTM 模型,具有较高的应用价值。

关键词:深度学习;CNN;LSTM;FCM;多传感器;未知复合故障;故障诊断

中图分类号:TP277

文献标志码:A

文章编号:1000-582X(2020)09-093-08

Fault diagnosis of multi-sensor signal with unknown composite fault based on deep learning

XING Liwen^a, YAO Wenkai^a, HUANG Ying^b

(a. Graduate School; b. Information and Communication, Engineering University of PAP, Xi'an 710086, P. R. China)

Abstract: Deep learning has been widely used in the field of fault diagnosis, among which convolution neural networks (CNN) and long short-term memory networks (LSTM) are typical models. As a common method of multi-sensor signal fault diagnosis, CNN can obtain good diagnosis effect, but it can not diagnose unknown composite fault. In order to solve this problem, a CNN-LSTM-FCM(fuzzy C-means) model was proposed. Since LSTM was more sensitive to the time signals with the connection between the front and the back, it was combined with CNN to realize the diagnosis of unknown signals. The decoupling of composite fault was achieved through the probability classification output. The CNN-LSTM-FCM model itself had optimized parameter design, which further improved the diagnosis accuracy. The chemical process fault measurement data was used for experiments, and the results showed that the diagnostic accuracy of CNN-LSTM-FCM model could reach 97.15%, which was superior to both CNN model and LSTM model, thus having a high application value in fault diagnosis.

Keywords: deep learning; CNN; LSTM; FCM; multi-sensor; unknown compound fault; fault diagnosis

收稿日期:2020-03-10

基金项目:国家自然科学基金资助项目(61573366)。

Supported by National Natural Science Foundation of China(61573366).

作者简介:邢砾文(1997—),女,硕士,主要从事故障诊断方向研究,(E-mail)544260972@qq.com。

现代装备功能繁多、结构复杂,同一装备中的不同系统之间相互交织、相互依赖。元件的故障会引起部件、系统的故障,系统之间的关联性也可能会造成故障的传播,使故障具有复合性。同时,实验环境下采集到的故障样本难以覆盖全部故障类型,使故障具有未知性,提高故障诊断难度。使用多个传感器进行信号采集,可以多角度、全方位检测装备运行状态,提高故障诊断效果。

传统故障诊断方法往往通过信号处理手段、专家经验、浅层学习等方式提取特征,特征的选择直接影响故障诊断效果,也难以充分利用海量数据。深度学习因其多隐层网络,能够自适应地从海量数据中提取特征,摆脱人为进行特征选择的局限性,其中,CNN 在处理海量数据上有独特优势,在轴承^[1-2]、旋转机械^[3-4]等领域得到应用。文献[1]将迁移学习与 CNN 相结合,在训练数据与实际数据有不同分布时也能较好诊断。文献[2]用复 Morlet 小波的连续小波变换处理时域信号,得到时频图像,再输入卷积神经网络,充分利用信号的时频特征。文献[3]提出多尺度深度 CNN 模型,拓宽、深化神经网络,达到更好的鲁棒性,缩短训练时间。文献[4]将多堆栈胶囊设计为解耦分类器,与 CNN 结合,准确解耦复合故障。吴魁^[5]等人将多传感器信号构建为测量数据帧,输入 CNN,完成了多传感器信号的融合。CNN 能够实现准确的故障诊断,但只能诊断已知、单一类别的故障,迫切需要其他方法解决未知、复合故障诊断问题。

LSTM 对具有前后联系的时间信号更加敏感,可以应用于故障诊断和预测,在锂电池^[6]、风冷式冷水机组系统^[7]、盾构机^[8]、轴承^[9]等领域得到应用。文献[6]改进 LSTM 结构,将输入门和遗忘门进行耦合,对新输入和历史单元状态的元素乘积进行筛选,筛选出更有益的信息;在输出门中加入来自“恒定误差旋转木马”的窥视孔连接,屏蔽了不需要的误差信号。文献[7]增加自适应增强模块,能够自适应地改变反向传播误差,并通过自适应设置学习率减少计算时间。文献[8]利用层规范化方法改进 LSTM,使网络的收敛速度更快、精度更高。文献[9]提出了一种改进具有全局性质的并行 LSTM。该模型由每个采样点的并行状态组成,通过加入窥视孔连接来提取序列的局部信息和全局信息,通过重用存储单元的信息更好提取序列信息。既充分利用信息,保证了迭代速度。虽然 LSTM 具有较好故障诊断效果,但是在大量数据的情况下,LSTM 网络将难以提取数据的非线性特征,且收敛速度慢。

CNN 与 LSTM 的组合已经在无人机^[10]、网络^[11]、单元设备^[12]、工业^[13]、轴承^[14-15]等领域得到了应用。该模型首先利用 CNN 提取特征,然后通过 LSTM 来探索时间序列数据的内在联系。其中 CNN 擅长数据处理,可以充分利用海量数据而不丢失数据之间的联系。将提取的特征向量输入 LSTM,有利于对时间序列数据的分析。该模型不仅保留了输入数据的特征及其相互关系,且保证模型的时序性。FCM 作为模糊分类方法,一个对象可以同时属于多个不同的类,其隶属度在 0~1 之间,能够应用于故障诊断,但诊断效果不如 CNN。

因此,研究将 LSTM 和 FCM 与 CNN 相结合,使其能够有效诊断未知信号,并通过概率分类输出实现复合故障的解耦,最终实现含未知复合故障的多传感器信号故障诊断。

1 关键技术

1.1 CNN 基本原理

典型的卷积神经网络 LeNet-5 的结构如图 1 所示。网络分为输入层、卷积层、子采样层、全连接层和输出层。

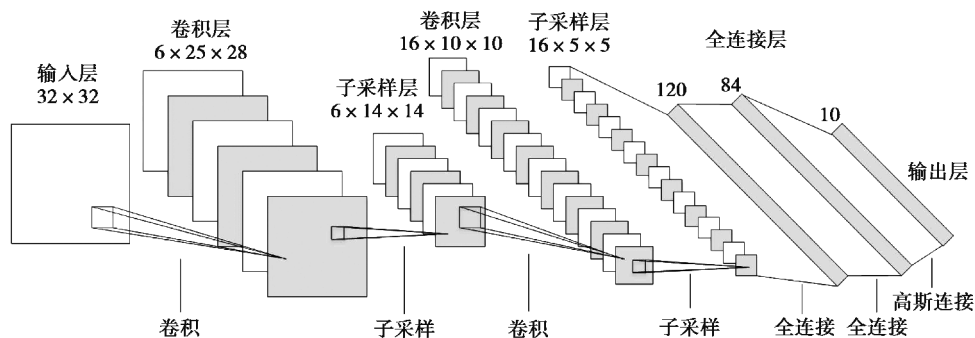


图 1 LeNet-5 结构

Fig. 1 The structure of LeNet-5

卷积层也称为特征提取层,卷积层通过不同的卷积核卷积前一层,每个卷积核对应一个提取特征,与前一层的感受连接。卷积核共享网络权重,每个卷积核提取有限的信息,因此通常使用多个卷积核来获取更多特征。

卷积运算的输出是

$$x = f(\sum x * w_{ij} + b), \tag{1}$$

其中: w_{ij} 是卷积核, b 是偏置, x 是输入特征图, $*$ 是卷积算子, $f(\cdot)$ 是激活函数。

子采样层也称为特征映射层,对从卷积层提取的特征进行操作。通过下采样操作,大大减少节点数量、网络连接数量和网络计算的复杂性。

下采样的输出是

$$x = f(\beta \text{down}(x) + b), \tag{2}$$

其中: β 是权值, b 是偏置, $\text{down}(\cdot)$ 是下采样函数, $f(\cdot)$ 是激活函数。

全连接层中的所有神经元节点都与前一层节点互连。它的输出是

$$h(x) = f(wx + b), \tag{3}$$

其中: x 是全连接层的输入, w 是权重, b 是偏置, $f(\cdot)$ 是激活函数。

1.2 LSTM 基本原理

LSTM 是循环神经网络(RNN, recurrent neural network)的一个变种,它解决了梯度反向传播过程中的梯度消失和爆炸问题。LSTM 的关键是 LSTM 单元,它可以决定是否维护状态信息。LSTM 单元结构如图 2 所示。

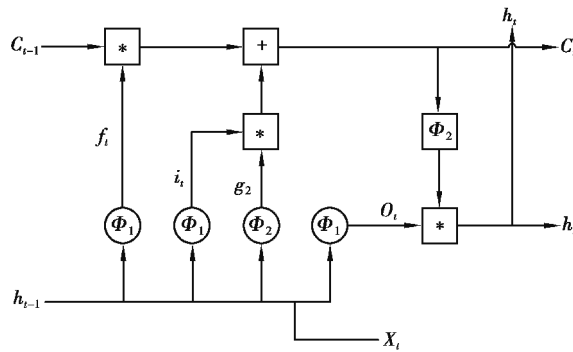


图 2 LSTM 单元结构

Fig. 2 The structure of LSTM Cell

LSTM 单元包含输入门、遗忘门和输出门。输入门用来决定输入信息,以保存单元的状态。遗忘门用于最后一次确定单元的状态。输出门决定 LSTM 单元的输出。输出的 c_t 和 h_t 反复输入到 LSTM 单元。具有正向传播的 LSTM 单元计算如下

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f), \tag{4}$$

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i), \tag{5}$$

$$g_2 = \tanh(W_c * [h_{t-1}, x_t] + b_c), \tag{6}$$

$$c_t = f_t * c_{t-1} + i_t * g_2, \tag{7}$$

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o), \tag{8}$$

$$h_t = o_t * \tanh(c_t), \tag{9}$$

公式(4),(5)和(6)分别是遗忘门、输入门和输出门。 W_f 、 W_i 和 W_c 分别是遗忘门、输入门和输出门的权重。 $\sigma(\cdot)$ 和 $\tanh(\cdot)$ 是 sigmoid 函数和双曲正切函数。

1.3 FCM 基本原理

FCM 算法由 Dunn 提出并由 Bezdek 进一步发展,是一种常用的聚类分析方法。FCM 将一组 s 维向量 $\mathbf{X} = (X_1, \dots, X_n)$ 划分为 c 个组,其中 $\mathbf{X}_j = (X_{j1}, \dots, X_{jm})'$ 表示 $j = 1, \dots, n$ 的第 j 个样本。第 i 组有中心向量

$\mathbf{v}_i = (v_{i1}, \dots, v_{is})$, FCM 的目的是确定聚类中心 \mathbf{v}_i , 其中“1ic”。对于第 j 个样本 \mathbf{X}_j 和第 i 个聚类中心 \mathbf{v}_i , $u_{ij} (\in [0, 1])$ 表示样本 \mathbf{X}_j 属于聚类中心 \mathbf{v}_i 的隶属度, 形成模糊隶属度矩阵 $\mathbf{U} = (u_{ij})_{n \times c}$ 。目标函数 J 的定义如下

$$\mathbf{v}_i = \frac{\sum_{j=1}^n u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^n u_{ij}^m}, \quad (10)$$

$$d_{ij} = \left(\sum_{k=1}^s (v_{ik} - x_{jk})^2 \right)^{\frac{1}{2}}, \quad (11)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}}, \quad i = 1, \dots, c; j = 1, \dots, n, \quad (12)$$

$$\min(J(\mathbf{U}, \mathbf{V})) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2, \quad (13)$$

对于 $\sum_{i=1}^c u_{ij} = 1, \forall j = 1, \dots, n$, (10) 和 (12) 中的 m 用于调整权重。

FCM 算法可描述如下:

1) 选择一个整数 c 和一个判别精度 ϵ , 在区间 $[0, 1]$ 内用 $c \times n$ 个随机数初始化模糊隶属度矩阵 \mathbf{U} 。

2) 根据式 (10) 计算 $\mathbf{v}_i (i = 1, \dots, c)$ 。

3) 分别根据式 (11) 和式 (12) 计算 d_{ij} 和 u_{ij} , 用新计算出的 u_{ij} 更新模糊隶属度矩阵 \mathbf{U} 。

4) 用式 (13) 计算目标函数 J 。如果 J 2 个相邻值之间的差小于判别精度 ϵ , 则停止计算。否则转到步骤 2)。

1.4 CNN-LSTM-FCM 基本原理

CNN-LSTM-FCM 模型如图 3 所示

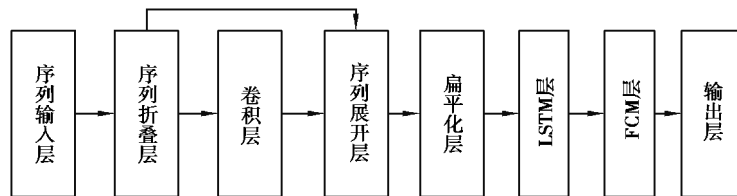


图 3 CNN-LSTM-FCM 模型结构

Fig. 3 The structure of CNN-LSTM-FCM model

模型通过 CNN 进行特征提取, 利用 LSTM 预测信号标签, 将二者组装为一个网络, 通过组合 2 个网络中的图层来直接对信号进行分类, 模型的输出为属于每一个类别的概率。

序列折叠层用于将输入的序列构建成测量数据帧。

卷积层对每组信号进行卷积, 实现特征提取。根据输入数据的维数, 卷积运算可分为一维卷积和二维卷积, 二维卷积通常用于处理图像。由于模型的输入是多个传感器信号, 一个传感器就是一维度信号, 多组传感器信号构成了一个矩阵, 因此使用的是一维卷积。

序列展开层和扁平化层用于恢复序列结构、输出矢量序列。

LSTM 层和输出层用于对输出的矢量序列进行分类。LSTM 层的输入信号中, 既包含特征信号, 又包含原始信号, 通过信号融合, 使输入信号中包含更多信息, 提高了诊断准确率。在 LSTM 层后设置了 FCM 层, 作为整个模型优化参数设计的一部分, 进一步提高了诊断精度。

对于任何分类模型来说, 只能对训练过的信号进行分类, 把未知信号归为已知类别。笔者构建的模型, 是用回归的思想解决含未知信号的分类问题。例如: 训练集包含五种类别, 在测试集中加入一种未知类别。测试时, 若输入 1 类信号, 那么输出在 1 左右; 若输入未知信号, 则输出在 0.5 左右, 不接近任何一类, 从而识别出未知信号。若输入信号中含有多种未知信号, 不能对它们进一步分类, 只能统一划归为未知信号。

传统 CNN 只能输出某一特定类的概率,而 CNN-LSTM-FCM 模型可以分别输出每一类的概率,实现复合故障的解耦。例如:有 5 组传感器信号,1 和 3 为故障信号,其余正常,将这组数据输入传统 CNN 模型,输出可能是故障 1,也可能是故障 3,但只能输出其中一种故障。若输入 CNN-LSTM-FCM 模型,则模型的概率输出为 $[1\ 0\ 1\ 0\ 0]$,从而诊断出同时具有故障 1 和 3。

模型的损失函数如下

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y_i = j\} \lg \frac{e^{\theta_j^T x_i}}{\sum_{l=1}^k e^{\theta_l^T x_i}} \right], \quad (14)$$

其中: x_i 为输入数据, y_i 为所属类别, m 为样本数, θ 为参数集,用于使损失函数最小化。

利用 CNN-LSTM-FCM 模型进行故障诊断的过程如下:

- 1) 按照文献[5]提出的方法构建测量数据帧并将其归一化。
- 2) 将数据集的 80% 构建为训练集,20% 为测试集,在测试集中加入未知故障信号。
- 3) 利用训练集对 CNN-LSTM-FCM 模型进行预训练。
- 4) 利用测试集测试模型故障诊断效果。

2 实验及分析

2.1 构建数据集

选用 mathworks 官网上发布的化学过程故障测量数据,共有 5 个传感器采集数据,每个传感器信号中均有正常信号和故障信号。对于同时段的 5 个传感器信号,每个信号截取 500 个采样点,按照图 4 和图 5 的方式构建 52×500 的测量数据帧并归一化。图中的 1~5 行分别表示传感器 1~5 的信号,每个传感器信号有 20% 的概率为故障信号,实现了数据的随机选择,不同传感器信号的故障是否相互独立,使得最终构建的测量数据帧包含各种故障类型的随机组合。将 80% 的数据集用作训练集,20% 为测试集。在训练集中,传感器 1~4 均包含正常信号和故障信号,传感器 5 个包含正常信号。在测试集中,加入传感器 5 的故障信号,作为未知故障。每个传感器信号的部分正常状态、故障状态分别如图 4、5 所示。

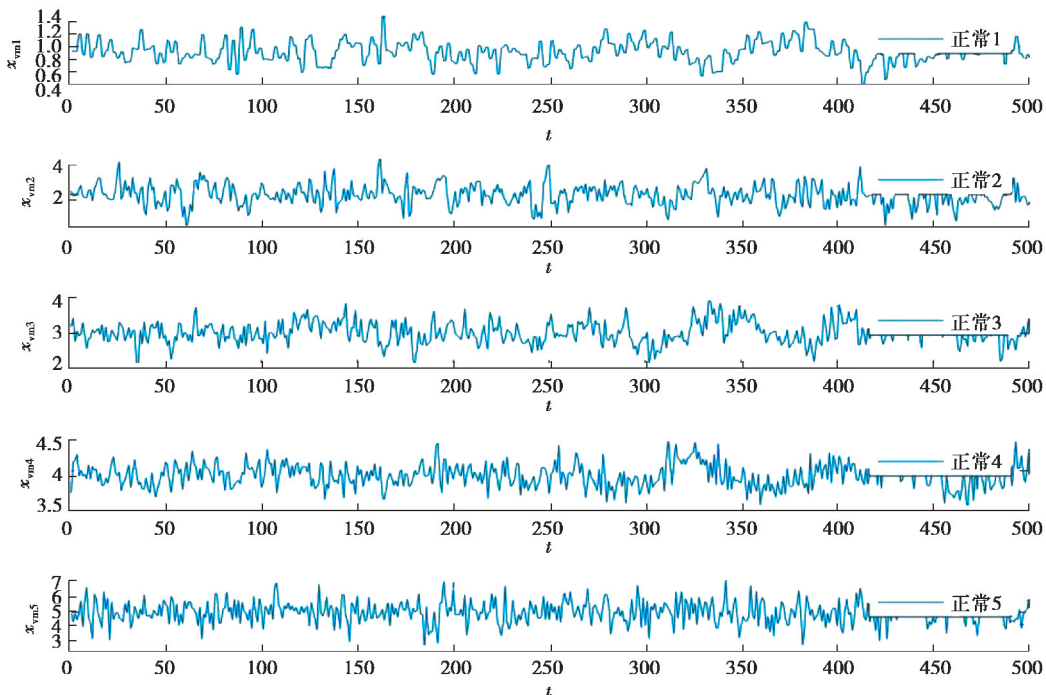


图 4 正常信号

Fig. 4 Normal signal

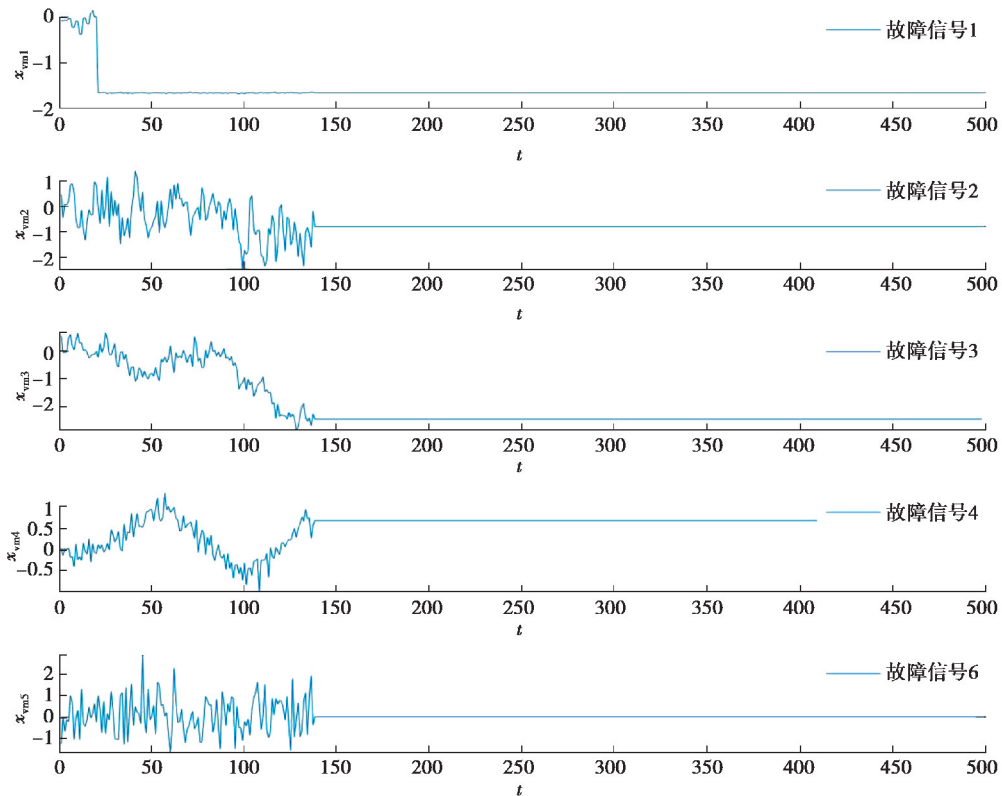


图 5 故障信号
Fig. 5 Fault signal

2.2 实验及分析

先用训练集对模型进行预训练,然后用测试集检测故障诊断效果。将提出的模型与 CNN、LSTM 进行比较,不同模型对应的损失值如图 6 所示,准确率如图 7 所示,数值如表 1 所示。从图、表所示的信息来看,CNN-LSTM-FCM 模型的损失值最低、准确率最高、收敛速度最快,其次是 LSTM 模型,CNN 模型诊断效果最不理想。实验结果表明,对于含未知复合故障的多传感器信号故障诊断,CNN-LSTM-FCM 模型具有更好的诊断效果,解决传统故障诊断模型只能诊断出已知、单一类别故障的问题。

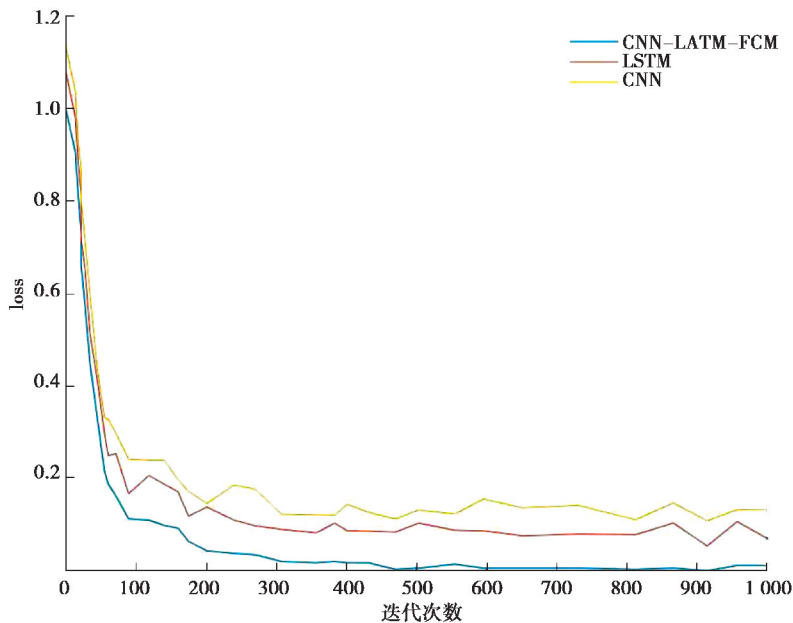


图 6 损失值
Fig. 6 Loss value

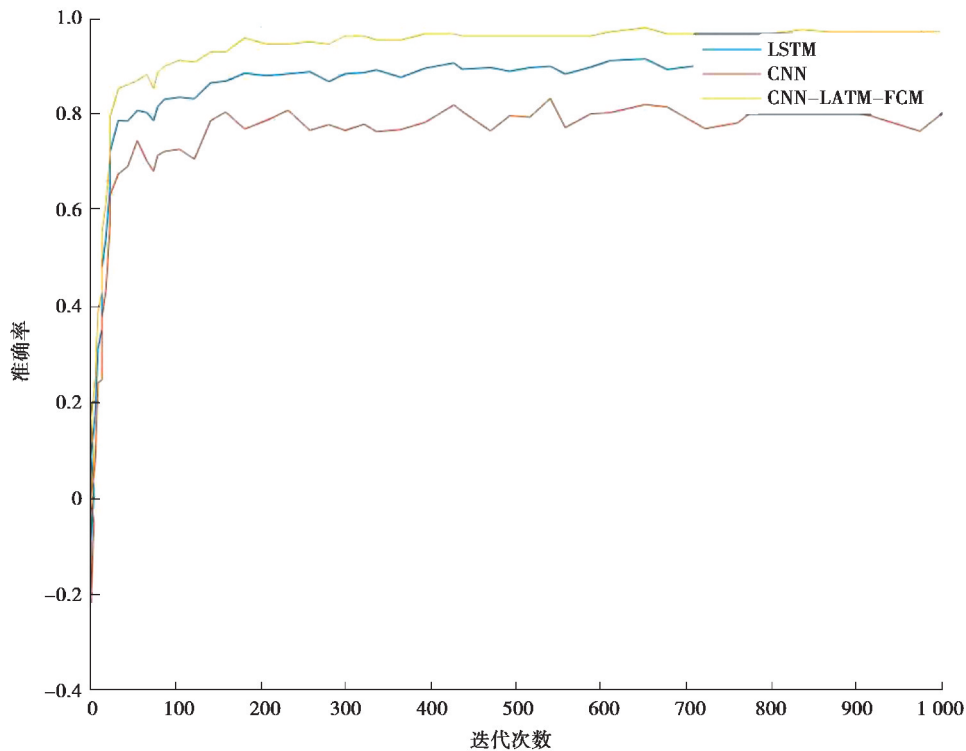


图 7 准确率

Fig. 7 Accuracy

表 1 各模型损失值、准确率

Table 1 Loss value and accuracy of each model

模型	损失值	准确率
CNN-LSTM-FCM	0.011 5	0.971 5
LSTM	0.070 0	0.891 9
CNN	0.132 9	0.800 4

3 结 论

随着装备功能不断完善,装备结构愈发复杂,往往需要用到多个传感器采集信号,其故障具有未知性、复合性,传统故障诊断模型只能诊断已知、单一类别故障,迫切需要其他方式解决这一问题。将 LSTM 和 FCM 与 CNN 进行组合,解决含未知复合故障的多传感器信号故障诊断问题,并通过与 CNN、LSTM 模型对比验证了其优越性。

提出的 CNN-LSTM-FCM 模型主要有以下优势:1)用回归的思想解决了含未知信号的分类问题。2)能够实现概率分类输出,解耦复合故障。3)通过优化参数设计提高了诊断准确率。

模型还有以下问题有待改进:1)层数和参数设置主要凭借经验,需要多次尝试才能达到较好诊断效果。2)使用传感器测量的数据,有时会因为传感器本身原因,造成测量数据存在问题,如何处理这种 nan 类型数据也有待研究。可以通过差值法、剔除法等方式解决这个问题,本实验选择的是剔除法。

参考文献:

[1] Guo L, Lei Y G, Xing S B, et al. Deep convolutional transfer learning network: a new method for intelligent fault diagnosis of machines with unlabeled data[J]. IEEE Transactions on Industrial Electronics, 2019, 66(9): 7316-7325.

- [2] Gao D W, Zhu Y S, Wang X, et al. A fault diagnosis method of rolling bearing based on complex morlet CWT and CNN [C/OL]. 2018 Prognostics and System Health Management Conference (PHM-Chongqing). Piscataway, NJ: IEEE, 2018 (2018-01-19)[2020-05-25]. <https://doi.org/10.1109/PHM-Chongqing.2018.00194>
- [3] Zhuang Z L, Qin W. Intelligent fault diagnosis of rolling bearing using one-dimensional multi-scale deep convolutional neural network based health state classification[C/OL]. 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC). Piscataway, NJ: IEEE, 2018 (2018-05-21) [2020-05-25]. <https://doi.org/10.1109/ICNSC.2018.8361296>
- [4] 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.
- [5] 吴魁, 王仙勇, 孙洁, 等. 基于深度卷积网络的多传感器信号故障诊断方法研究[J]. 计算机测量与控制, 2018, 26(1): 18-21.
WU Kui, WANG Xianyong, SUN Jie, et al. Study of multi-sensor fault diagnose method based on convolutional neural networks[J]. Computer Measurement & Control, 2018, 26(1): 18-21. (in Chinese)
- [6] Li P H, Zhang Z J, Xiong Q Y, et al. State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network[J]. Journal of Power Sources, 2020, 459: 228069.
- [7] Gao L, Li D H, Li D, et al. An improved LSTM based sensor fault diagnosis strategy for the air-cooled chiller system[C/OL]. 2019 Chinese Control Conference (CCC). Piscataway, NJ: IEEE, 2019(2019-10-17)[2020-05-25]. <https://doi.org/10.23919/ChiCC.2019.8866339>
- [8] Sun Z H, Zheng H Y, Shi B H. Faults prediction and diagnoses of shield machine based on LSTM[C/OL]. 2019 Chinese Control And Decision Conference (CCDC). Piscataway, NJ: IEEE, 2019(2019-09-12)[2020-05-25]. <https://doi.org/10.1109/CCDC.2019.8832341>
- [9] Zhang Z Q, Zhou F N, Chen D M. Application of improved parallel LSTM in bearing fault diagnosis[C/OL]. 2019 Chinese Automation Congress (CAC). Piscataway, NJ: IEEE, 2019(2020-04-13)[2020-05-25]. <https://doi.org/10.1109/CAC48633.2019.8997417>
- [10] Fu J, Sun C, Yu Z, et al. A hybrid CNN-LSTM model based actuator fault diagnosis for six-rotor UAVs[C/OL]. 2019 Chinese Control And Decision Conference (CCDC). Piscataway, NJ: IEEE, 2019(2019-09-12)[2020-05-25]. <https://doi.org/10.1109/CCDC.2019.8832706> 2019.
- [11] Tan Z L, Pan P S. Network fault prediction based on CNN-LSTM hybrid neural network[C/OL]. 2019 International Conference on Communications, Information System and Computer Engineering (CISCE). Piscataway, NJ: IEEE, 2019 (2019-08-22)[2020-05-25]. <https://doi.org/10.1109/CISCE.2019.00113>
- [12] Zheng L, Xue W H, Chen F, et al. A fault prediction of equipment based on CNN-LSTM network[C/OL]. 2019 IEEE International Conference on Energy Internet (ICEI). Piscataway, NJ: IEEE, 2019(2019-08-08)[2020-05-25]. <https://doi.org/10.1109/ICEI.2019.00101>
- [13] Gao Y, Gong P, Li L X. An end-to-end model based on CNN-LSTM for industrial fault diagnosis and prognosis[C/OL]. 2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC). Piscataway, NJ: IEEE, 2018 (2018-11-08)[2020-05-25]. <https://doi.org/10.1109/ICNIDC.2018.8525759>
- [14] Du W, Wang Y X. Stacked convolutional LSTM models for prognosis of bearing performance degradation[C/OL]. 2019 Prognostics and System Health Management Conference (PHM-Qingdao). Piscataway, NJ: IEEE, 2019(2019-12-26) [2020-05-25]. <https://ieeexplore.ieee.org/document/8943063>
- [15] Qiao M Y, Yan S H, Tang X X, et al. Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads[J]. IEEE Access, 2020, 8: 66257-66269.