

DOI:10.19651/j.cnki.emt.2108237

## 基于深度学习与注意力机制的化工故障分类\*

唐颖川<sup>1</sup> 黄姣茹<sup>1</sup> 钱富才<sup>1,2</sup>

(1. 西安工业大学电子信息工程学院 西安 710021; 2. 西安理工大学陕西省复杂系统控制与智能信息处理重点实验室 西安 710048)

**摘要:** 针对现有的故障诊断方法在处理高维度且动态特征明显的化工生产过程中观测的数据时,存在无法识别长时间依赖关系、精确度不够的问题,对长短时记忆模型进行改进,提出了一种基于深度学习与 attention 机制的分类模型,以田纳西-伊斯曼仿真平台的仿真数据作为研究对象,通过小波阈值去噪法对数据进行预处理,再对模型分类效果进行验证,比较了本文模型与改进前的模型,最后通过 t-snc 算法绘制样本数据及在模型各层输出特征向量在二维空间的分布图。实验结果表明,改进后的深度学习模型,对故障分类时能达到 92.71% 的召回率与 93.05% 的准确率,相对改进前的模型分别提高了 16.84% 与 13.66%,对数据特征的学习效果更好,更适用于化工数据。

**关键词:** 深度学习;attention 机制;故障诊断

**中图分类号:** TP391.5 **文献标识码:** A **国家标准学科分类代码:** 520.2060

## Chemical fault classification based on deep learning and attention mechanism

Tang Yingchuan<sup>1</sup> Huang Jiaoru<sup>1</sup> Qian Fucui<sup>1,2</sup>

(1. School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China; 2. Key Laboratory of Shanxi Province for Complex System Control and Intelligent Information Processing, Xi'an University of Technology, Xi'an 710048, China)

**Abstract:** In view of the problems that the existing fault diagnosis methods cannot identify the long-time dependence relationship and have insufficient precision when processing the observed data in the chemical production process with high dimensions and obvious dynamic characteristics, the long-term memory model is improved, and a classification model based on depth learning and attention mechanism is proposed. Then the model classification effect is verified and the model is compared with the model before improvement. Finally, the sample data is drawn and the distribution of feature vectors in two-dimensional space is output at each level of the model by t-sne algorithm. The experimental results show that the improved in-depth learning model can achieve a recall rate of 92.71% and an accuracy rate of 93.05% for fault classification, which are improved by 16.84% and 13.66% respectively compared with the model before the improvement. The learning effect on data characteristics is better, and it is more suitable for chemical data.

**Keywords:** deep learning;attention mechanism;fault diagnosis

## 0 引言

随着化工工业的迅速发展,化工生产的规模越来越大,大规模的生产使得企业的效益成倍提升<sup>[1]</sup>,但复杂的生产过程也使得出现生产故障的可能性变的更高。因此实时监控过程数据,及时发现异常情况并且判断出故障类型很重要。

化工生产过程监控的数据具有动态、高维度的特性,可

测的过程变量和故障之间的关系复杂<sup>[2]</sup>。最初数据驱动的分类方法以统计学和机器学习为主<sup>[3-4]</sup>,如文献[5]结合了支持向量机与 PCA 的方法用于 TE 过程数据的故障分类;文献[6]提出了一种基于核主元分析与关联向量机(RVM)相结合的故障诊断算法(KPCA-RVM)。近年来计算机的性能得到提升,在处理大数据时深度学习的优势逐渐展现。深度学习中能够通过算法本身提取特征值,在应用于难以人工提取特征的复杂数据时具有优势。文献[7]通过结合

收稿日期:2021-11-01

\* 基金项目:陕西省重点研发计划一般项目(2021GY-067)资助

数据增强与深度卷积网络(CNN)的方法对 TE 过程仿真数据进行了研究,得到了不错的结果。但是 CNN 由于结构的局限性,只能对数据的静态特征进行训练,忽略了序列数据间在时间上长期的依赖关系。循环神经网络相比传统方法增加了循环结构解决了长时间依赖关系,在此基础上文献[8]采用了长短时记忆网络(long short term memory, LSTM),增加了记忆单元来选择性的记忆过去的信息,缓解了循环结构在处理长序列问题的不足,因此近年来成为了工业故障诊断的热门,如文献[9]提出了一种基于长短期记忆(LSTM)的方法直接对原始过程数据进行分类,实验结果在精确度上较于 RNN 有所提高。虽然 LSTM 在处理长序列数据时具有结构上的优势,但在面对高维度的数据时,仍然存在不足。高维数据各变量对于故障的判断而言存在相关度的差异,模型能够关注更重要的变量时,分类的精度必然更高。为了使算法具有识别变量相关性程度的能力,一些学者采用了数据统计的方法对数据降维,去掉了不相关的变量。如文献[10]提出了一种基于组合核函数 KPCA 与改进 ELM 的故障诊断方法,但是这种方法在降维的同时会丢失很多局部信息,影响诊断结果。

attention 机制的提出提供了一种在不丢失数据的前提下量化各特征权重的方法,文献[11]结合了卷积网络与 attention 机制,在图片分类的任务中提高了对人脸微小运动的识别能力,这是因为注意力机制增加关键特征的权重后,能够使模型忽略干扰信息,识别样本间的微小差异。在本文的研究中将改进 LSTM 模型,并将其与 attention 机制结合,使得模型在兼具对数据的局部感知能力与对动态特征识别能力的前提下,能够识别故障数据间微小的差异,提高模型对高维度序列数据的识别精度。

## 1 Tennessee Eastman 化工过程

### 1.1 TE 过程仿真原理

在实际中化工数据大多是不能向外公开的,得到真实的采样数据对于研究者来说是一件困难的事情,而化工生产的研究又非常必要,为了研究人员能够较容易的获得实验数据,美国 Eastman 化学公司开发了 Tennessee Eastman(TE)仿真平台<sup>[12]</sup>,是根据伊斯曼化学公司的实际工艺做少许修改后得到的。

TE 过程由反应器、冷凝器、压缩机、气液分离器和汽提塔组成,在工业过程中操作 A、C、D、E 四种反应物在这些单元中发生化学反应。化学反应如式(1)和(2)所示。



G 与 H 是所需的产物,此外生产过程中还发生了两种副反应,如式(3)和(4)所示。



其中, F 为反应的副产物。为了确保反应的正常进行,

还需要在反应器中加入惰性气体 B。反应器中的产物会送至冷凝器进行冷却,再送至气液分离器中分离出所需的液体 G 与 H。再将剩余的反应物通过汽提塔与循环压缩机送回反应器,完成循环。

### 1.2 数据预处理

本文将数据集分为训练集以及测试集两部分,其中数据集由 TE 仿真平台 48 h 运行下获得,观测数据总数为 960,每个样本由 52 个过程变量组成,其中故障在 8 h 时引入,前 160 个观察值为正常数据,中间 600 个值作为训练集,后 200 个值作为测试集。本文将训练模型对其中的 6 种故障进行识别,用以验证训练的效率和分类的精度。

由于化工过程生产环境复杂操作复杂等因素<sup>[13]</sup>,监测的数据会受到各种噪声的影响,这些噪声会影响对过程数据的分析和数据特征学习。本文将采用小波降噪算法对 TE 过程数据降噪<sup>[14]</sup>。

小波阈值去噪方法主要由 3 个步骤组成,如图 1 所示。

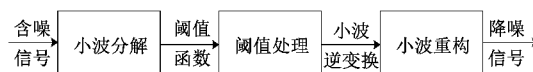


图 1 小波阈值去噪过程

根据数据的特点,选择了 Daubechies 小波作为基小波,并且采用了一种改进的自适应分解层数的阈值方法,如式(5)所示。

$$\lambda_j = \sigma \sqrt{2 \ln N_j} / (\sqrt{2})^{j-1} \quad (5)$$

式中:  $\sigma$  为估计的噪声标准方差;  $N_j$  为第  $j$  层分解的系数个数;  $j$  为小波分解尺度。由于深度学习模型具有很强的鲁棒性,能够适用输入数据的震荡,本文选择能更好保留输入信息的硬阈值函数处理方法,其表达式如式(6)所示。

$$\hat{W}_{j,k} = \begin{cases} W_{j,k}, & |W_{j,k}| \geq \lambda \\ 0, & |W_{j,k}| < \lambda \end{cases} \quad (6)$$

式中:  $W_{j,k}$  为含噪信号的小波系数;  $\hat{W}_{j,k}$  为经过阈值处理后的小波系数;  $j$  为分解尺度;  $k$  为系数序号;  $\lambda$  为小波阈值。

## 2 CNN-LSTM-Attention 模型

### 2.1 模型的优点

为了应对化工过程数据动态高维度的特点,本文采用 CNN-LSTM 的深度学习结构<sup>[15,16]</sup>,结合了 attention 机制的特点<sup>[17]</sup>,增加了相关性强的变量的权重,使得在训练模型时能够更加关注与对故障分类更有决定性的变量,从而提高训练的效率和分类的精度。其结构如图 2 所示。

本文提出的模型具有如下优点:

1) 采用循环神经网络作为框架的核心部分,能有效地关注到数据中的动态特征。

2) LSTM 由于结构的特点感受野会受到限制,通过与

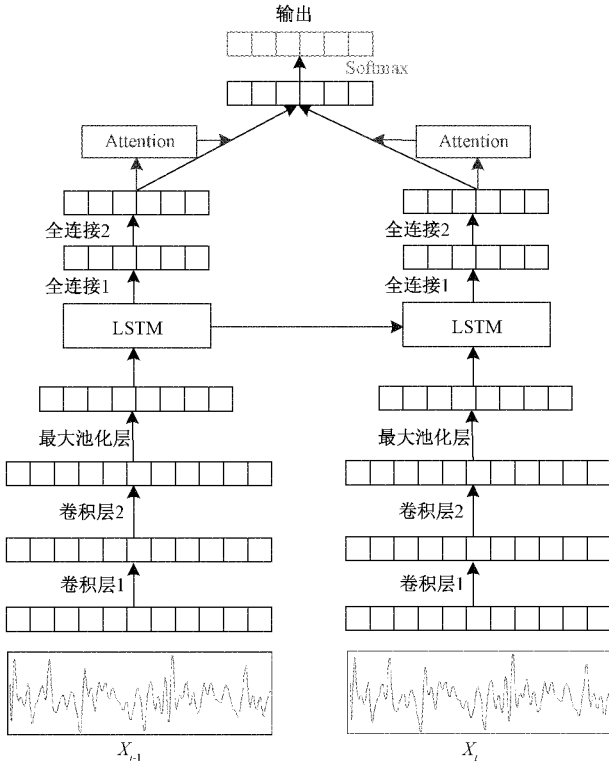


图 2 CNN-LSTM-Attention 模型

CNN 的结合能够提高整个深度学习框架的感受野,使得每次训练接受的输入信息更加全面,从而提升深度学习的训练效率。

3)在模型的输出层前增加了 Attention 结构,使得模型在训练的过程中能够更关注相关性高的特征。

### 2.2 模型的组成

本文模型由卷积层、最大池化层、循环层、全连接层、以及 Softmax 和输出层组成,在全连接层与 Softmax 层之间加入 attention 机制,用来筛选最重要的特征值,模型输出的结果与真实结果进行对比,计算交叉熵损失函数,如式(7)所示。

$$Loss = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (7)$$

式中:  $n$  代表分类的总数,  $y$  代表期望的输出,  $a$  代表神经元的实际输出,  $x$  表示第几种类别。

Loss 函数的大小代表着模型判断结果与真实分类情况的差异,在训练过程中,通过改变深度学习中各层的权重,减小 Loss 函数的大小,就是在优化模型的分​​类准确度。

#### 1)一维卷积层

数据集数据有较长的时间维度,直接传入 LSTM 网络会面临梯度爆炸的问题,通过卷积层初步提取特征再传入循环层的方式可以有效地减少循环层训练的参数,并且提高训练的效率。在一维卷积网络中,卷积层会在感受野中逐层提取输入序列数据的特征,通过迭代选择出最有效的高级特征,其结构如图 3 所示。

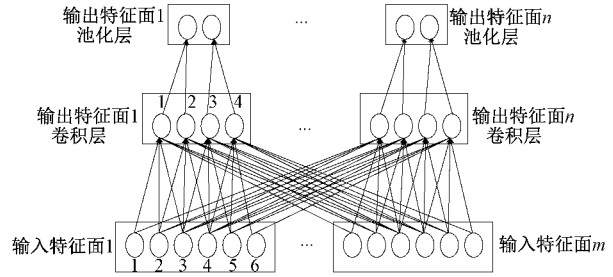


图 3 一维卷积网络

可见,较高层中的神经元通过上一层的部分神经元加权传递信息,这部分神经元就是高层神经元的感受野,通过多层的传递后,输出层神经元的感受野足够大,包含了整体数据的信息,从而做到了权值共享,减小了模型的复杂度,便于网络的训练。

#### 2)循环层

数据集具有动态特征,循环神经网络可以有效地提取此类特征,通过对卷积层提取特征进行训练,可以得到包含时序信息的更有效的特征值。LSTM 中的记忆单元由如图 4 所示,主要由 3 个部分组成,分别是输入门、遗忘门、输出门以及神经单元。

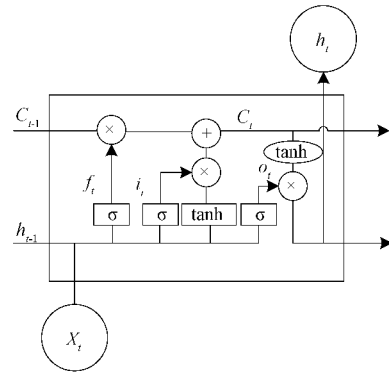


图 4 LSTM 记忆单元

在前向传递的过程中,记忆单元的输入包括当前时刻的数据  $x_t$  和上一时刻的输出  $h_{t-1}$ 、cell 单元状态  $C_{t-1}$ 。序列数据  $x(t)$  和前一时刻的输出  $h(t-1)$  组成输入数据,先进入遗忘门,在经过激活函数处理后得到  $f(t)$ ,其表达式如式(8)所示。

$$f_t = \sigma(\mathbf{W}_f [\mathbf{h}_{t-1}, \mathbf{X}_t]) + \mathbf{b}_f \quad (8)$$

其中,  $\mathbf{W}_f$  与  $\mathbf{b}_f$  为遗忘门的权重与偏置,经过训练后可以得到准确值,且不唯一,  $\sigma$  为 sigmoid 激活函数。遗忘门的输出  $f(t)$  取值范围在  $[0, 1]$ ,表示传递上一层神经元中  $c(t-1)$  的概率。之后  $x(t)$  与  $h(t-1)$  进入输入门层,先是经过 sigmoid 激活函数,得到  $i(t)$ ,此数值将决定当前的神经元状态  $c(t)$  的更新情况,数据再通过 tanh 激活函数得到临时神经元状态  $\bar{C}_t$ ,经过  $i(t)$  的处理后传入  $c(t)$  中,与遗忘门共同决定神经元状态,其表达式如式(9)~(11)所示。

$$i_t = \sigma(\mathbf{W}_i [\mathbf{h}_{t-1}, \mathbf{X}_t] + \mathbf{b}_i) \quad (9)$$

$$\bar{c}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{X}_t] + \mathbf{b}_c) \quad (10)$$

$$\mathbf{c}_t = f_t \circ \mathbf{c}_{t-1} + \mathbf{i}_t \circ \bar{\mathbf{c}}_t \quad (11)$$

其中,  $\mathbf{W}_i$ 、 $\mathbf{b}_i$ 、 $\mathbf{W}_c$ 、 $\mathbf{b}_c$  为输入门的权重与偏置;“ $\circ$ ”为向量按元素乘法。最后数据进入输出门,输出门决定当前的输出状态  $h(t)$ ,由两部分组成,首先数据通过一个 sigmoid 激活函数计算出  $o(t)$ ,同时更新后的神经元状态函数  $c(t)$  通过一个 tanh 激活函数与  $o(t)$  共同计算出当前输出状态  $h(t)$ 。其表达式如式(12)与(13)所示。

$$\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{X}_t] + \mathbf{b}_o) \quad (12)$$

$$\mathbf{h}_t = \mathbf{o}_t \circ \tanh(\mathbf{c}_t) \quad (13)$$

其中,  $\mathbf{W}_o$ 、 $\mathbf{b}_o$  为输出门权重与偏置。

### 3) attention 机制

通过卷积层与循环层的训练后,可获得有效的特征,但是对于识别某一故障而言,特征之间是存在主次之分的,并且各故障由于产生的原因以及表现的形式不同,依赖的主要特征也不同,attention 机制可以量化这种对不同特征的依赖关系。在深度学习中,attention 机制是模仿人类视觉的注意力,使得模型能够更关注与研究对象关系密切的特征。

如图 5 所示,在 attention 机制中,通过权值的方式对各个特征进行了处理,故障类型和各特征的相关程度可以通过计算得出。各特征对应 Value 的权值系数进行加权求和,得到最终的 Attention 数值。attention 机制通过给各特征值加权的方式,将特征值与故障类型的相关性进行了量化。

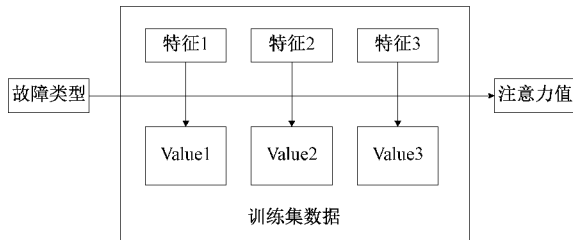


图 5 注意力机制

## 3 仿真实验及结果分析

本文将在模型训练时的学习效率、模型的泛化性两个方面对提出深度学习模型的性能进行分析。通过与 LSTM 模型进行比较,从对故障分类的精度以及模型训练的效率,验证本文提出方法的效果。训练效率由在训练集中损失函数的迭代速度来体现,模型的泛化性由对测试集数据的分类精度来验证。流程如图 6 所示。

首先为了排除噪声对验证结果造成干扰,通过小波阈值函数法对训练数据预处理,部分数据处理结果如图 7 所示。

可见通过预处理可以过滤到大部分噪声,将处理后的数据作为数据集,输入诊断模型中进行验证实验。分别对 LSTM 与 CNN-LSTM-Attention 模型进行测试,由 python

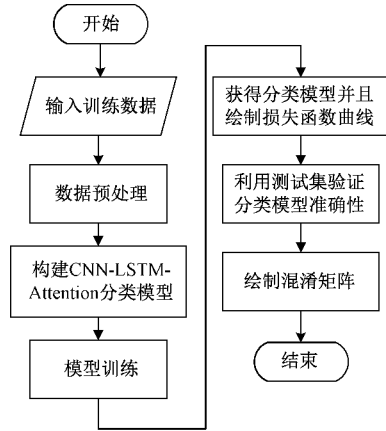


图 6 训练流程

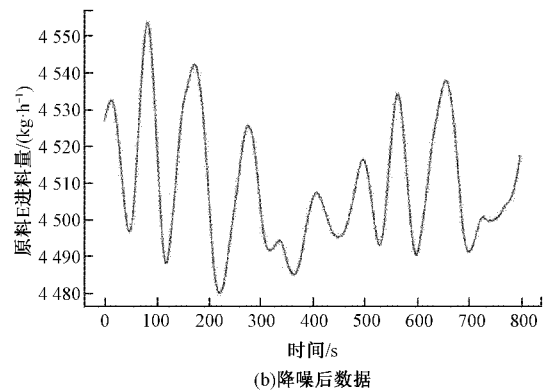
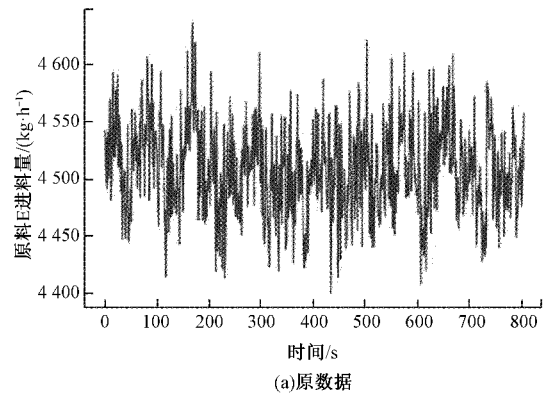


图 7 小波降噪结果

通过训练过程的回调函数绘制混淆矩阵可视化结果,如图 8 所示。

其中,行为数据的真实标签,列为模型预测标签,处于对角线上的色块代表正确的分类,其上的数字表示占测试样本的百分比,具体的数值如表 1 与 2 所示。

表 1 为 LSTM 模型对测试集进行分类的混淆矩阵,其中行代表测试数据的真实值,列代表模型的分类情况,召回率指对目标样本而言,被正确预测的概率。准确度指对预测结果而言,预测为正的样本有多少是真正的正样本。可以看出,在面对高维度的时序数据时,LSTM 对各类故障的平均召回率只有 0.7587,准确度也仅有 0.7939,模型能准



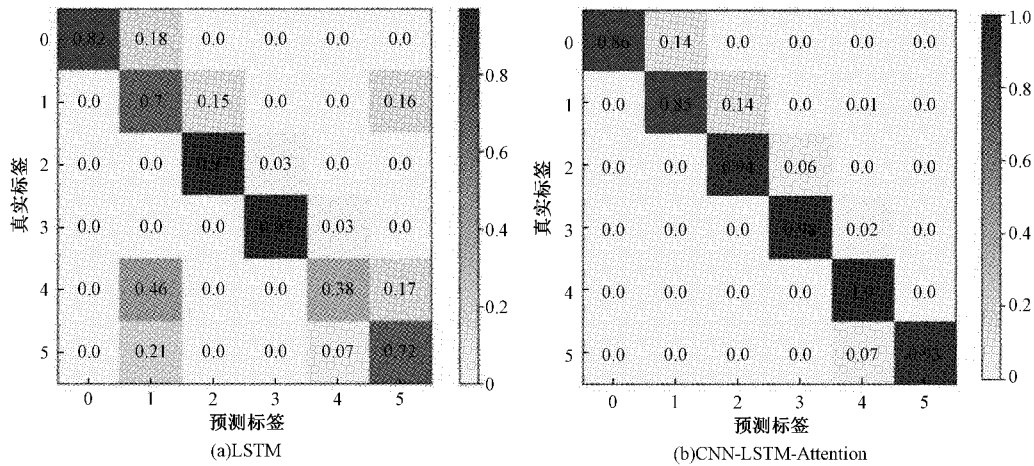


图 8 LSTM 与 CNN-LSTM-Attention 的混淆矩阵热力图

表 1 LSTM 测试结果混淆矩阵

预测标签	真实标签						样本数	召回率
	0	1	2	3	4	5		
0	79	17	0	0	0	0	96	0.822 9
1	0	67	14	0	0	15	96	0.697 9
2	0	0	93	3	0	0	96	0.968 8
3	0	0	0	93	3	0	96	0.968 8
4	0	44	0	0	36	16	96	0.375 0
5	0	20	0	0	7	69	96	0.718 8
预测数	79	148	107	96	46	100	576	0.758 7
精确度	1.0	0.452 7	0.869 2	0.968 8	0.782 6	0.690 0	0.793 9	

表 2 CNN-LSTM-Attention 测试结果混淆矩阵

预测标签	真实标签						样本数	召回率
	0	1	2	3	4	5		
0	83	13	0	0	0	0	96	0.864 6
1	0	82	13	0	1	0	96	0.854 2
2	0	0	90	6	0	0	96	0.937 5
3	0	0	0	94	2	0	96	0.979 2
4	0	0	0	0	96	0	96	1
5	0	0	0	0	7	89	96	0.927 1
预测数	83	95	103	100	106	89	576	0.927 1
精确度	1	0.863 2	0.873 8	0.940 0	0.905 7	1	0.930 5	

确的识别部分故障,但是对一些故障的识别能力却不足,可见改进前的 LSTM 在分类时序数据时是存在局限性的。

表 2 为 CNN-LSTM-Attention 模型对测试集进行分类的混淆矩阵,其中平均召回率为 0.927 1,平均准确率为 0.930 5。对比 LSTM 训练结果,可以看出 CNN-LSTM-Attention 在处理高维度大数据时能在测试数据中有更良好的表现,证明模型具有更强的泛化能力。模型在训练后针对各故障都学习到了有效的特征,由此可见 CNN-LSTM-Attention 不

仅在训练过程中具有高效的特征提取能力,在结合大数据时具有优势,并且在分类精度上也相比 LSTM 模型有大幅度的提高。

图 9 为通过 python 记录的训练过程中, Loss 函数的变化情况,对比可知, CNN-LSTM-Attention 模型通过 50 次的训练后, Loss 函数快速的收敛,并且在之后的训练中保持着高的分类精度,而 LSTM 模型 Loss 函数收敛的较慢,并且会由于梯度爆炸或者消失出现较大的波动,证明

其训练过程中无法有效地提取时序数据间的长时间依赖关系。可以看出本文的模型在处理动态高维度数据时能更有效的学习到各类故障的特征。

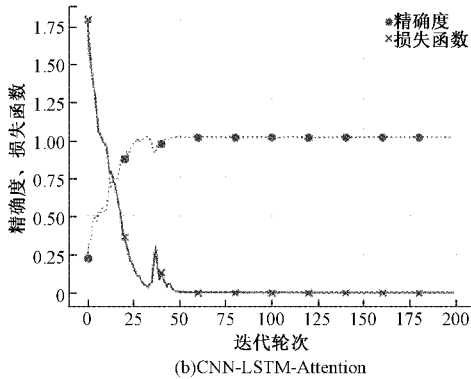
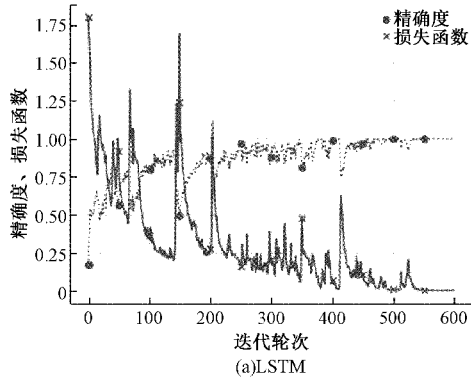


图9 LSTM与CNN-LSTM-Attention  
Loss函数与分类精确度

为了分析测试数据在CNN-LSTM-Attention模型各层的分布变化,本文通过t-sne(t-distributed stochastic neighbor embedding)的方法将包含52个变量的测试样本数据,以及各层输出的特征向量投影到2维空间中,再绘制其空间分布图,通过观察数据在模型中的空间变化过程,证明故障诊断模型的有效性。原始数据及各层特征向量的空间分布图如图10所示。

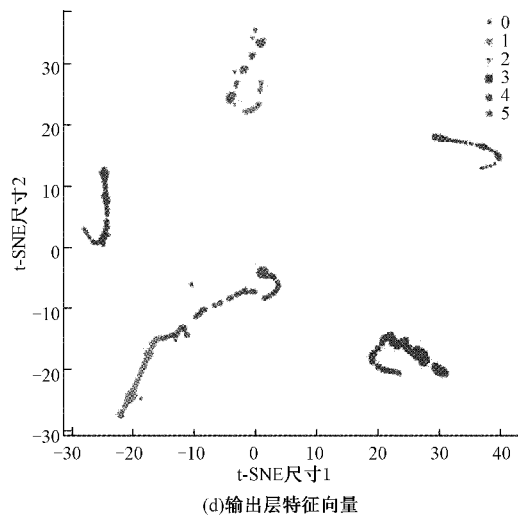
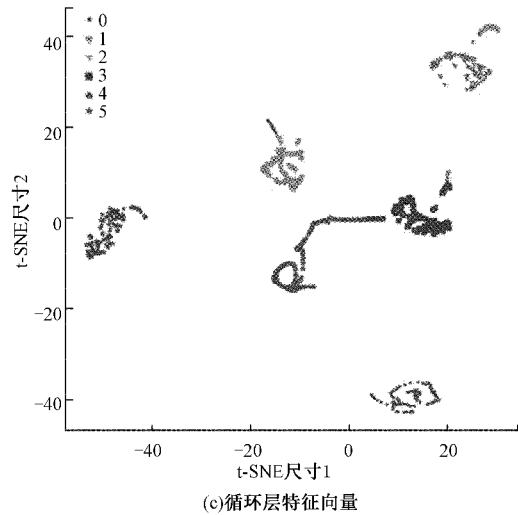
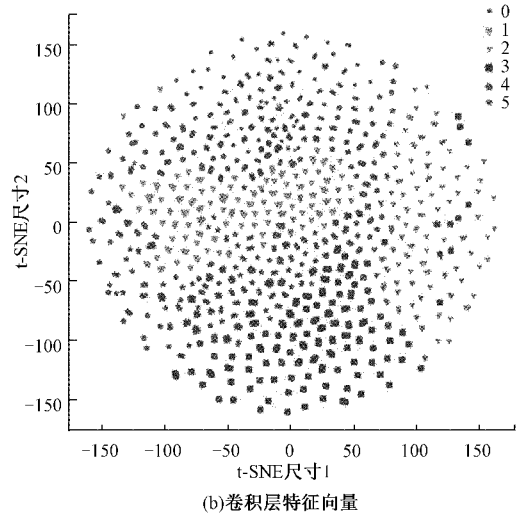
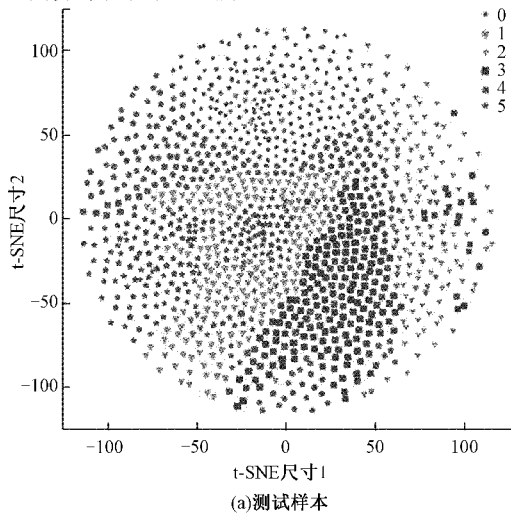


图10 测试样本数据及其在模型各层输出的特征  
向量2维空间分布图

图10中6种形状的点各代表一类故障样本在2维空间中的投影,可见图10(a)中原始样本在空间中混淆在了一起,难以区别;图10(b)表示经过卷积层的空间变换以及

池化处理输出特征向量的空间分布情况,可见相比原始数据相同颜色点在空间中的分布更加集中;图 10(c)表示经过循环层的空间变换后输出的特征向量,可见同类间的距离更大了;图 10(d)表示模型输出层的特征向量,可见原始样本经过整个模型的处理后,按照不同的分类被清晰的分成了 6 个部分。通过可视化空间变换的过程可以证明,本文提出的模型能有效地提取出高维度时间序列数据中的特征,并且完成分类。

#### 4 结 论

循环神经网络在处理时间序列数据时具有优势,通过循环结构可以关注到序列数据的长期依赖关系,其变体在解决梯度爆炸等问题后更是推动了该类深度学习算法在工业生产过程中的应用。本文为了提高模型的训练效率与泛化性,采用了将深度学习与注意力机制相结合的 CNN-LSTM-Attention 深度学习算法,在仿真数据集中表现出了高效的训练过程以及准确的测试结果,这体现出了本文算法的在处理动态高维度数据时的优势。现实生产中故障数据的数量远小于正常数据的数量,这会给模型识别故障数据带来问题,如何处理数据的不平衡性是深度学习分类模型的又一挑战方向。

#### 参考文献

- [1] 刘强,秦泗钊. 过程工业大数据建模研究展望[J]. 自动化学报, 2016, 42(2): 3-13.
- [2] 文成林,吕菲亚,包哲静,等. 基于数据驱动的微小故障诊断方法综述[J]. 自动化学报, 2016, 42(9): 1285-1299.
- [3] 文成林,吕菲亚. 基于深度学习的故障诊断方法综述[J]. 电子与信息学报, 2020, 42(1): 234-248.
- [4] 孟晨,杨华晖,王成,等. 数据驱动的武器系统电子元部件级故障诊断研究综述[J]. 系统工程与电子技术, 2021, 43(2): 574-583.
- [5] CHEN J, XIN G, ZHU X, et al. Fault classification on tennessee eastman process: PCA and SVM[C]. International Conference on Mechatronics & Control, Jinzhou: IEEE, 2015, 2194-2197.
- [6] 赵小强,薛永飞,杨武. 基于 RVM 的改进 KPCA 化工过程故障诊断算法研究[C]. 第 32 届中国控制会议论文集(D 卷), 2013. 1181-1185.
- [7] GU X, SONG H, WANG T, et al. Chemical process fault diagnosis based on mixup-convolution neural network[C]. 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), Yangzhou: IEEE, 2019, DOI: 10.1109/ICEMI46757. 2019. 9101484.
- [8] 谭峰,李成南,萧红,等. 基于 LSTM 循环神经网络的数控机床热误差预测方法[J]. 仪器仪表学报, 2020, 41(9): 79-87.
- [9] ZHAO H, SUN S, BO J. Scquential fault diagnosis based on LSTM neural network[J]. IEEE Access, 2018, 6(99): 12929-12939.
- [10] 钱锬. 基于组合 KPCA 与改进 ELM 的工业过程故障诊断研究[D]. 重庆:重庆大学, 2016.
- [11] WANG F, TAX D. Survey on the attention based RNN model and its applications in computer vision[J]. ArXiv, 2016, ArXiv: 1601. 06823.
- [12] 王培良,夏春江. 基于 PCA-PDBNs 的故障检测与自主学习辨识[J]. 仪器仪表学报, 2015, 36(5): 1147-1154.
- [13] 王康成,尚超,柯文思,等. 化工过程深度神经网络软测量的结构与参数自动调整方法[J]. 化工学报, 2018, 69(3): 900-906, 1253.
- [14] 齐晓辉,宋宛鸿,李文召. 一种改进的小波去噪方法在 OTDR 中的应用[J]. 光通信技术, 2021, 45(4): 54-58.
- [15] DONAHUE J, HENDRICKS L A, ROHRBACH M, et al. Long-term recurrent convolutional networks for visual recognition and description [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(4): 677-691.
- [16] 程焕新,成凯,蒋泽芹. 基于注意力机制的 CNN 人脸表情识别[J]. 电子测量技术, 2021, 44(10): 128-132.
- [17] 李春,张代红,李登武. 基于 CNN-LSTM 的配电网供电能力研究[J]. 国外电子测量技术, 2019, 38(9): 16-21.

#### 作者简介

唐颖川, 硕士, 主要研究方向为数据驱动等。

E-mail: t\_oblivious@163.com

黄姣茹, 博士, 主要研究方向为系统识别等。

E-mail: 465194985@qq.com

钱富才, 教授, 主要研究方向为复杂系统建模与优化等。

E-mail: fcqian@xaut.edu.cn