基于MSE改进BiLSTM网络算法的工业互联网异常流量时空融合检测

程 光[1,2]，李沛霖[1]

（1. 北京联合大学 北京市信息服务工程重点实验室，北京 100101；2. 北京联合大学 前沿智能技术研究院，北京 100101）

摘 要：针对设备在通信传输过程中产生的大量数据信息容易成为黑客和恶意用户的攻击目标，从而产生异常流量，以及流量数据的稀疏性导致难以捕捉到全局特征之间的关联性，进而影响异常流量检测效果的问题，提出了基于均方误差（MSE）改进双向长短时记忆（BiLSTM）神经网络算法的工业互联网异常流量时空融合检测方法。首先，通过One-Hot编码将工业互联网流量数据转换为数值型数据，利用MSE中的挤压激励（SE）机制调整流量特征的权重，捕捉全局特征之间的关联性。其次，采用BiLSTM神经网络的正向与反向LSTM提取网络流量的时空融合特征。最后，将时空融合特征输入Softmax分类器中，对流量进行识别，实现异常检测。实验结果表明，当迭代次数达到30次时，本文方法的损失值可以降至0.4以下，当迭代次数达到60次时，F_1与马修斯相关系数均可达到60，证明该方法具有良好的整体性能。

关键词：多头挤压激励机制；BiLSTM神经网络；特征融合；异常流量检测；softmax分类器

中图分类号：TP393 **文献标志码**：A **文章编号**：1671-5497（2025）04-1406-06

DOI：10.13229/j.cnki.jdxbgxb.20240279

Spatio temporal fusion detection of abnormal traffic in industrial Internet based on MSE improved BiLSTM network algorithm

CHENG Guang[1,2]，LI Pei-lin[1]

（1. *Beijing Key Laboratory of Information Service Engineering*，*Beijing Union University*，*Beijing* 100101，*China*；2. *Frontier Intelligent Technology Research Institute*，*Beijing Union University*，*Beijing* 100101，*China*）

Abstract：Addressing the issue that the large amounts of data generated by devices during communication transmission are prone to becoming targets for hackers and malicious users，thereby generating abnormal traffic，and that the sparsity of traffic data makes it difficult to capture the associations between global features，which in turn affects the detection effectiveness of abnormal traffic，a spatiotemporal fusion detection method for abnormal traffic in industrial Internet of Things（IoT）based on the improved bidirectional long short-term memory（BiLSTM）neural network algorithm using mean squared error（MSE）is proposed. Firstly，the industrial Internet traffic data is converted into numerical data through the One-Hot coding method，and the SE mechanism in MSE is used to adjust the weight of traffic

收稿日期：2024-03-22.

基金项目：国家重点研发计划项目（2021YFB1715700）.

作者简介：程光（1964-），男，教授，博士. 研究方向：先进制造和工业工程. E-mail：chengguang@buu.edu.cn

characteristics to capture the correlation between global characteristics. Secondly, using the forward and backward LSTM of BiLSTM neural network, the spatiotemporal fusion features of network traffic are extracted. Lastly, and the spatio temporal fusion features are input into the softmax classifier to identify traffic and achieve anomaly detection. The experimental results show that when the number of iterations reaches 30, the loss value of the proposed method can reach below 0.4, when the number of iterations reaches 60, both F1 and Matthews correlation coefficients can reach 60, proving that this method has good overall performance.

Key words: multi head squeezing incentive mechanism; BiLSTM neural network; feature fusion; abnormal traffic detection; Softmax classifier

0　引　言

在工业互联网中,设备、传感器、控制系统、生产数据等都通过网络连接和通信,实现了设备之间的互联互通。在此背景下,海量数据流的汇聚,为企业日常管理与生产提供依据,但同时也造成了一定网络安全问题[1]。在网络攻击中,异常流量属于常见的手段,可能会导致网络拥塞、数据丢失和通信延迟,从而影响生产过程的正常进行。为了提高工业互联网的安全运行,需要检测网络中的异常流量,及时发现工业互联网中存在的潜在威胁,以此作出相关决策与预警。

文献[2]方法对网络流量数据的特征维度进行层次聚类划分,根据数据特征相似性距离将相关性较高的特征划分到同一特征子集中。针对每个特征子集,利用自动编码器进行特征约简,利用多层极限学习机构建异常流量检测模型,对约简后的特征进行检测。但网络流量数据存在缺失值、异常值或噪声,需要进行有效的数据清洗和处理,以确保模型训练的有效性和准确性。文献[3]方法使用滑动窗口划分网络流量,对划分后的流量开展小波变换,引入链式稀疏自编码器(Sparse auto-encoder,SAE)映射处理变换后的序列完成序列重构,并将重构结果输入分类器中,获得初步检测结果,通过加权投票策略对多个初步检测结果进行融合处理,获得最终检测结果。但该方法在空间映射过程中存在较大误差,序列重构精度还需进一步提升。文献[4]方法对网络流量数据进行了预处理,以提高数据分布的均衡性。该结合结合门控循环单元与聚合残差变换网络构建特征提取模块,将网络流量数据输入其中以提取流量特征,利用Softmax分类器完成异常检测。但该方法无法有效地控制Softmax分类器的学习率,网络的收敛效果还需提升。文献[5]方法提取了网络流量特征,结合SMO与Taylor级数提出Taylor-SMO算法对特征进行分类处理,以此实现异常流量检测。但该方法提取的流量特征较片面,没有考虑流量的时间特性与空间特性,导致特征提取精度低。

为弥补上述研究的不足,本文采用多头挤压激励(Squeeze and excitation,SE)机制均方误差(Mean squared, MSE)改进(Bidirectional long short-time memory,BiLSTM)神经网络模型,建立用于异常流量检测的MSE-BiLSTM模型。通过多头SE机制MSE提取工业互联网流量的局部平行特征,将上述特征输入BiLSTM模型中,获得工业互联网流量的时空融合特征,最后通过Softmas分类器实现异常流量检测。

1　工业互联网异常流量时空融合检测

1.1　工业互联网流量数值型特征转换

通过以下3个步骤对工业互联网流量进行预处理,将其转换为数值型特征:

(1)通过One-Hot编码预处理工业互联网流量。

(2)标准化处理工业互联网流量r_{ij},将其转换为数值型数据:

$$\begin{cases} r'_{ij}=(r_{ij}-A_j)/S_j \\ A_j=(r_{1j}+r_{2j}+\cdots+r_{nj}) \\ S_j=(|r_{1j}-A_j|+\cdots+|r_{nj}-A_j|)/n \end{cases} \tag{1}$$

式中:r'_{ij}为标准化处理后的数值型数据;A_j为数值型数据r_{ij}的平均值;S_j为数值型数据r_{ij}的平均绝对偏差;n为数据量。

(3)通过下式归一化处理r'_{ij},使其处于[0,1]范围内:

$$r''_{ij}=(r'_{ij}-r_{\min})/(r_{\max}-r_{\min}) \tag{2}$$

式中:$r_{\max}$、$r_{\min}$分别为r'_{ij}的最大值和最小值。

1.2 基于MSE的异常流量特征权重校准方法

卷积神经网络(Convolutional neural networks, CNN)具有良好的特征提取能力,且可有效采集目标的特征,被广泛应用于一维序列数据与二维图像数据的处理中[6,7]。利用一维CNN可分析一维序列数据,一维CNN主要由一维卷积层和池化层部分构成:一维卷积层可以用来获取样本数据中隐藏的局部特征;池化层通过二次采样处理在保留关键信息的同时实现样本降维。

通过交替连接池化层与卷积层可构成具备语义信息学习能力的一维CNN。调查发现,网络流量数据属于一维序列数据,包含通信事件与网络连接情况,因此直接通过一维CNN对其进行处理,可获得其局部模式与特征关联[8,9]。

CNN主要用于分析数据之间的局部关联性,但由于流量数据本身具有稀疏性,在分析过程中容易出现语义信息丢失的现象,难以捕捉全局特征之间的关联,降低了异常检测结果的精度。而MSE中的SE机制能够调整流量特征的权重,帮助捕捉全局特征的关联性,从而提高异常检测过程中对全局特征的感知能力。为提高CNN在工业互联网异常流量检测过程中的感知能力,通过MSE对流量特征权重进行校准处理,MSE结构如图1所示。

由图1可知,MSE属于具有强可移植性的轻量级模型,模型中包含多个并行的SE机制,其主要作用是调整网络流量的特征权重。将归一后的数值型特征输入模型中,通过SE模块中的挤压操作G_{sq}生成全局统计信息z_c:

$$z_c=r''_{ij}G_{sq}(x_c)=r''_{ij}\sum_{c=1}^{H}x_c/H \tag{3}$$

式中:x_c为尺寸为H的第c个特征。利用SE模块中的激励操作G_{ex}获得z_c的加权系数d:

$$d=G_{ex}[G_{sq}(x_c),W] \tag{4}$$

式中:W为全连接层对应的权重。多头SE机制可有效融合不同视角信息,以此提高特征潜在依赖的挖掘能力。基于MSE改进BiLSTM网络算法的工业互联网异常流量时空融合检测方法通过多个SE模块对流量的特征进行并行拟合学习,将各SE模块生成的$d^i(i=1,2,3)$与原始流量特征相乘,获得加权处理后的流量特征,采用Concat对上述特征进行融合处理,获得MSE的输出X':

$$X'=z_c\mathrm{Concat}(d^1X,d^2X,d^3X) \tag{5}$$

1.3 基于BiLSTM层的流量时空融合特征提取

BiLSTM是在LSTM基础上变换而来的网络结构,由正向LSTM和反向LSTM构成[10,11],分别用于提取网络流量的正向特征与反向特征,具体结构如图2所示。BiLSTM通过这种双向处理机制可以有效结合未来与过去的上下文信息,挖掘流量之间存在的潜在关联,提高特征提取的全面性[12,13]。最后,将MSE的输出X'输入BiLSTM层中,获得工业互联网流量的时空融合特征。BiLSTM的输出在每个时间步上由以下几个部分控制:

(1)遗忘门f_t:负责BiLSTM网络的遗忘操

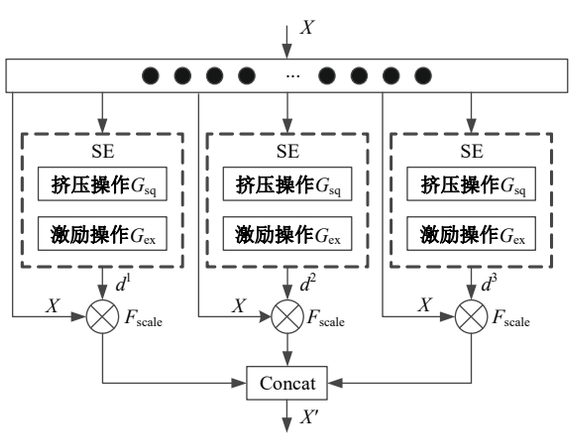

图1 MSE结构

Fig. 1 MSE structure

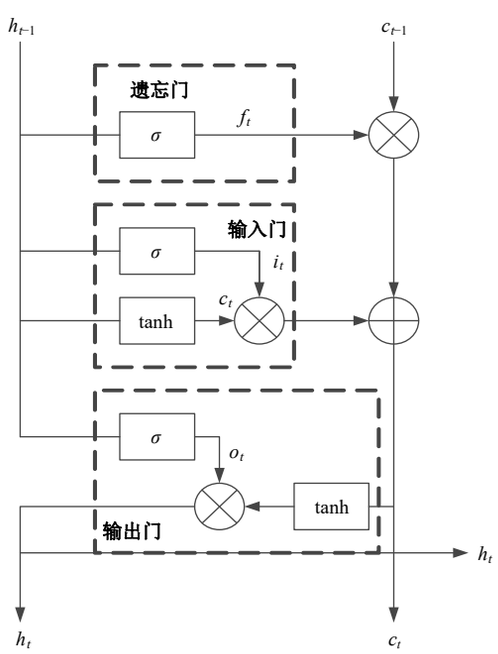

图2 BiLSTM结构

Fig. 2 BiLSTM structure

作,该层的权重与偏置分别为w_f、b_f。

(2)细胞状态更新c_t:负责信息的存储,判断当前细胞状态下需要存储哪些信息。

(3)输出门o_t:选择并控制输出的信息,该层的权重与偏置分别为w_o、b_o;。

(4)输入门i_t:选择并控制输入的信息,该层的权重与偏置分别为w_i、b_i。

在BiLSTM模块中通过正向与反向LSTM操作提取工业互联网流量MSE输出X'的特征。

正向LSTM操作可表示为:

$$\begin{cases} i_t=\sigma\{w_i\cdot[h_{t-1},X']+b_i\} \\ f_t=\sigma\{w_f\cdot[h_{t-1},X']+b_f\} \\ o_t=\sigma\{w_o\cdot[h_{t-1},X']+b_o\} \\ c_t=f_t*c_{t-1}+i_t*\tilde{c}_t \end{cases} \tag{6}$$

式中:$\sigma\{\}$为激活函数;*为元素乘运算;h_{t-1}为LSTM的实际输出;$\tilde{c}_t$为待选值向量,可通过tanh函数获得;c_{t-1}为细胞上层状态。

反向LSTM操作可表示为:

$$\begin{cases} i_t=\sigma\{w_i\cdot[h_{t+1},X']+b_i\} \\ f_t=\sigma\{w_f\cdot[h_{t+1},X']+b_f\} \\ o_t=\sigma\{w_o\cdot[h_{t+1},X']+b_o\} \\ c_t=f_t*c_{t+1}+i_t*\tilde{c}_t \end{cases} \tag{7}$$

由此,融合正向LSTM与反向LSTM获得工业互联网流量的时空融合特征X_t=[正向LSTM,反向LSTM]。

1.4　基于Softmax的异常流量检测

Softmax分类器是一种多类别分类器,主要用于将神经网络最后一层的输出转换为对应各个类别的概率分布。Softmax分类器可以对模型输出的特征进行分类,计算输入数据属于各个类别的概率,从而实现对多类别问题的分类任务。为此,选用Softmax[14,15]作为MSE-BiLSTM模型的分类器,将BiLSTM层输出的时空融合特征X_t输入Softmax分类器中,最终获得工业互联网异常流量的检测结果:

$$P(y^{(i)}=j|x^{(i)})=X_t\frac{e^{\vartheta x^{(i)}}}{\sum_{j=1}^{K}e^{\vartheta x^{(i)}}} \tag{8}$$

式中:$P(y^{(i)}=j|x^{(i)})$为第i个流量特征样本属于j类的概率;K为流量类别数量;ϑ为学习因子。

设置Softmax分类器的损失函数$L(\vartheta)$,对其进行优化,通过调整Softmax分类器的学习率,提高模型的收敛速度:

$$L(\vartheta)=-\frac{1}{m}\left(\sum_{i=1}^{m}\sum_{j=1}^{K}1|y^{(i)}=j|\log\frac{e^{\vartheta x^{(i)}}}{\sum_{j=1}^{K}e^{\vartheta x^{(i)}}}\right)+\frac{\lambda}{2}\sum_{i=1}^{K}\sum_{j=0}^{n}\vartheta \tag{9}$$

式中:λ为权重衰减项;m为样本数量。

2　实验与分析

为验证本文方法的整体有效性,需要对进行测试。本次测试的实验环境与相关配置如表1所示。本次测试所用的数据集如表2所示。

表1 实验环境与配置

Table 1　Experimental environment and configuration

实验环境	配　置
Python	3.7.6
操作系统	Win10 LTSC2019
Tensorflow	2.0.0
处理器	Intel Core i7-6800K CPU 3.40 GHz
Keras	2.3.1
内存	16GB PDRP 2666 MHz
显卡	GTX 1080Ti

表2 测试数据集

Table 2　Test dataset

数据集	数据量
U2R	228
Normal	60 560
R2L	16 185
DoS	229 855
Probe	4 165

MSE模型中,设置SE模块数量为3个,激励机制数量为2,优化器为Adam,学习率为0.001;BiLSTM模型中,设置隐藏单元数为128,Dropout率为0.2,学习率为0.01,批量大小为32。

2.1　本文方法的收敛能力测试

选取单CNN模型、BiLSTM模型与本文方法的MSE-BiLSTM模型在表2数据集中进行训练,测试3种方法的损失值,测试结果如图3所示。

由图3可知,随着迭代次数的增加,MSE-BiLSTM模型、单CNN模型与BiLSTM模型的损失值均有所下降,但MSE-BiLSTM模型的损失值始终低于其他2种模型,在迭代次数达到30次时可以降至0.4以下,表明MSE-BiLSTM模型具有良好的收敛效果。

2.2　异常流量时空融合检测对比测试

将文献[3]方法(基于多尺度特征的网络流量异常检测方法)和文献[4]方法(基于时空融合深度学习的工业互联网异常流量检测方法)作为对

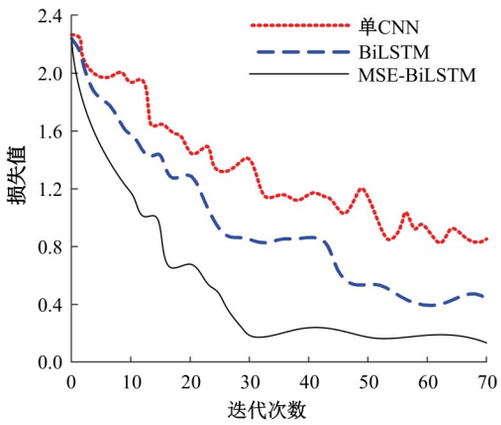

图3 不同模型的loss值

Fig. 3 Loss values of different models

比方法,与本文方法在相同测试环境下进行工业互联网异常流量检测测试,选取F_1值和马修斯相关系数MCC作为评估指标:

$$F_1=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{10}$$

$$\text{MCC}=\frac{\text{TP}\times\text{TN}-\text{FP}\times\text{FN}}{\sqrt{(\text{TP}+\text{FN})\times(\text{TN}+\text{FP})\times(\text{TP}+\text{FP})\times(\text{TN}+\text{FN})}} \tag{11}$$

式中:Recall为召回率;Precision为精确率;TP为网络中存在的异常流量被正确检测的数量;TN为正常流量被正确检测的数量;FP为正常流量被错误检测为异常流量的数量;FN为异常流量被错误检测为正常流量的数量。

基于式(10)(11),获取不同方法的工业互联网异常流量检测结果,如图4所示。分析图4可知,当迭代次数达到60次时,本文方法在异常流量检测过程中的F_1与MCC均可达到60,高于文献[3]方法与文献[4]方法。这是因为本文方法将MSE引入BiLSTM网络中,提高了流量特征的提取精度;并设置了损失函数,调整了Softmax分类器的学习率,从而提高了模型的收敛能力,进而提高了异常流量的检测精度。

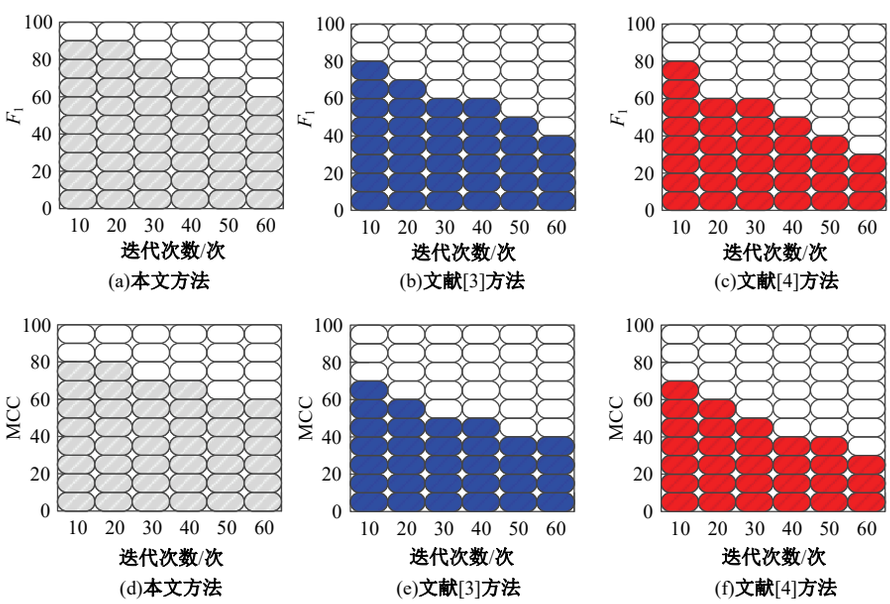

图4 异常流量检测结果

Fig. 4 Abnormal traffic detection results

3 结束语

针对目前工业互联网异常流量检测方法存在的问题,本文建立了MSE-BiLSTM模型。该模型具有优异特征学习能力,根据提取的流量特征,采用Softmax分类器实现异常流量检测。经验证,本文方法建立的模型在异常流量检测过程中表现出良好的收敛效果与较高的检测精度。

参考文献:

[1] 林广朋, 李闯. 入侵攻击下无线网络安全态势感知算法[J]. 计算机仿真, 2023, 40(12): 451-454, 547.
Lin Guang-peng, Li Chuang. Wireless network security situation awareness algorithm under intrusion attacks[J]. Computer Simulation, 2023, 40(12): 451-454, 547.

[2] 丁建立, 刘亦舟, 梁婷婷. 基于特征约简与多层极

限学习机的网络流量异常检测[J]. 现代电子技术, 2022, 45(5): 84-89.
Ding Jian-li, Liu Yi-zhou, Liang Ting-ting. Network traffic anomaly detection based on feature reduction and multi-layer extreme learning machine[J]. Modern Electronics Technique, 2022, 45(5): 84-89.

[3] 段雪源, 付钰, 王坤, 等. 基于多尺度特征的网络流量异常检测方法[J]. 通信学报, 2022, 43(10): 65-76.
Duan Xue-yuan, Fu Yu, Wang Kun, et al. Network traffic anomaly detection method based on multi-scale characteristic[J]. Journal on Communications, 2022, 43(10): 65-76.

[4] 胡向东, 张婷. 基于时空融合深度学习的工业互联网异常流量检测方法[J]. 重庆邮电大学学报: 自然科学版, 2022, 34(6): 1056-1064.
Hu Xiang-dong, Zhang Ting. Abnormal traffic detection method for industrial Internet based on deep learning with time-space fusion[J]. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition), 2022, 34(6): 1056-1064.

[5] Bhor H N, Kalla M. TRUST-based features for detecting the intruders in the Internet of Things network using deep learning[J]. Computational Intelligence, 2022, 38(2): 438-462.

[6] 宋建辉, 王思宇, 刘砚菊, 等. 基于改进FFRCNN网络的无人机地面小目标检测算法[J]. 电光与控制, 2022, 29(7): 69-73, 80.
Song Jian-hui, Wang Si-yu, Liu Yan-ju, et al. Ground small target detection algorithm of UAV based on improved FFRCNN network[J]. Electronics Optics and Control, 2022, 29(7): 69-73,80.

[7] 周云, 赵瑜, 郝官旺, 等. 基于应变信号时频分析与CNN网络的车辆荷载识别方法[J]. 湖南大学学报: 自然科学版, 2022, 49(1): 21-32.
Zhou Yun, Zhao Yu, Hao Guan-wang, et al. Vehicle load identification method based on time frequency analysis of strain signal and convolutional neural network[J]. Journal of Hunan University(Natural Science Edition), 2022, 49(1): 21-32.

[8] 程汪刘, 任仰勋, 倪修峰, 等. 基于改进Cascade R-CNN网络模型的防振锤缺陷识别[J]. 安徽大学学报:自然科学版, 2022, 46(5): 64-70.
Cheng Wang-liu, Ren Yang-xun, Ni Xiu-feng, et al. Defect recognition of vibration dampers based on improved Cascade R-CNN network model[J]. Journal of Anhui University(Natural Science Edition), 2022, 46(5): 64-70.

[9] 王得道, 王森荣, 林超, 等. 基于CNN-LSTM融合神经网络的CRTSⅡ型轨道板温度预测方法[J]. 铁道学报, 2023, 45(2): 108-115.
Wang De-dao, Wang Sen-rong, Lin Chao, et al. CRTS Ⅱ track slab temperature forecasting method based on CNN-LSTM fusion neural network[J]. Journal of the China Railway Society, 2023, 45(2): 108-115.

[10] 张浩, 胡昌华, 杜党波, 等. 多状态影响下基于Bi-LSTM网络的锂电池剩余寿命预测方法[J]. 电子学报, 2022, 50(3): 619-624.
Zhang Hao, Hu Chang-hua, Du Dang-bo, et al. Remaining useful life prediction method of lithium-ion battery based on Bi-LSTM network under multi-state influence[J]. Acta Electronica Sinica, 2022, 50(3): 619-624.

[11] 刘继, 顾凤云. 基于BERT与BiLSTM混合方法的网络舆情非平衡文本情感分析[J]. 情报杂志, 2022, 41(4): 104-110.
Liu Ji, Gu Feng-yun. Unbalanced text sentiment analysis of network public opinion based on BERT and BiLSTM hybrid method[J]. Journal of Intelligence, 2022, 41(4): 104-110.

[12] 罗晶, 高永, 梁葆华, 等. 基于CNN-BiLSTM网络模型的无人机飞行质量评价[J]. 工程数学学报, 2023, 40(2): 171-189.
Luo Jing, Gao Yong, Liang Bao-hua, et al. UAV flight quality evaluation based on CNN-BiLSTM Network model[J]. Chinese Journal of Engineering Mathematics, 2023, 40(2): 171-189.

[13] 王继东, 杜冲. 基于Attention-BiLSTM神经网络和气象数据修正的短期负荷预测模型[J]. 电力自动化设备, 2022, 42(4): 172-177, 224.
Wang Ji-dong, Du Chong. Short-term load prediction model based on Attention-BiLSTM neural network and meteorological data correction[J]. Electric Power Automation Equipment, 2022, 42(4): 172-177, 224.

[14] 封强, 潘保芝, 韩立国. 基于卷积降噪自编码器和Softmax回归的微地震定位方法[J]. 地球物理学报, 2023, 66(7): 3076-3085.
Feng Qiang, Pan Bao-zhi, Han Li-guo. Microseismic source location method based on convolutional denoising auto-encoder and Softmax regression[J]. Chinese Journal of Geophysics, 2023, 66(7): 3076-3085.

[15] 冯治广, 董佳佳, 王茂英. 基于Softmax的采摘机器人目标识别技术研究[J]. 农机化研究, 2023, 45(2): 184-188.
Feng Zhi-guang, Dong Jia-jia, Wang Mao-ying. Research on target recognition technology of picking robot based on Softmax[J]. Journal of Agricultural Mechanization Research, 2023, 45(2): 184-188.