

DOI:10.19322/j.cnki.issn.1006-4710.2024.02.011

<https://xuebao.xaut.edu.cn>

引文格式:王赛,张立新,陈乃源,阚希,王军昂,吴凯枫. 基于 CNN-BiLSTM-SA 网络的人类活动识别[J]. 西安理工大学学报, 2024, 40(2):253-259,290.

WANG Sai, ZHANG Lixin, CHEN Naiyuan, KAN Xi, WANG Junang, WU Kaifeng. Human activity identification based on CNN-BiLSTM-SA network[J]. Journal of Xi'an University of Technology, 2024, 40(2): 253-259,290.

基于 CNN-BiLSTM-SA 网络的人类活动识别

王 赛¹, 张立新^{1,2}, 陈乃源¹, 阚 希², 王军昂¹, 吴凯枫¹

(1. 南京信息工程大学 自动化学院, 江苏 南京 210044; 2. 无锡学院, 江苏 无锡 214105)

摘要: 针对传统的神经网络对人类活动行为识别精度不高的问题, 本文提出了一种基于双通道机制的卷积神经网络叠加双向长短期记忆网络和自注意力的混合网络模型(convolutional neural network-bi-directional long short-term memory-self-attention, CNN-BiLSTM-SA)。首先将数据集中的加速度和角速度数据作为网络的两个输入, 然后使用卷积神经网络叠加双向长短期记忆网络的模式搭建系统, 最后引入自注意力机制增强系统的分类能力。实验结果表明, 在 UCI-HAR 数据集中, 本网络的平均 F_1 分数达到 98.6%, 平均准确率达到 98.4%, 比卷积神经网络叠加长短期记忆神经网络模型(convolutional neural network-long short-term memory, CNN-LSTM)收敛速度更快并且准确率提高了 4.2%, 在人类活动识别方面, 具有更广阔的应用前景。

关键词: 人类活动识别; 传感器; CNN-BiLSTM; 自注意力机制

中图分类号: TP212 **文献标志码:** A **文章编号:** 1006-4710(2024)02-0253-07

Human activity identification based on CNN-BiLSTM-SA network

WANG Sai¹, ZHANG Lixin^{1,2}, CHEN Naiyuan¹, KAN Xi², WANG Junang¹, WU Kaifeng¹

(1. School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044, China;
2. Wuxi University, Wuxi 214105, China)

Abstract: In view of the problem that the traditional neural network is not accurate in recognizing human activities, this paper proposes a hybrid network model based on the two-channel mechanism of convolutional neural network superimposed with bi-directional long short-term memory network and self-attention (CNN-BiLSTM-SA). First, the acceleration and angular velocity data in the data set are used as the two inputs of the network, and then the system is built by using the convolution neural network to overlay the bidirectional short-term and short-term memory network; finally, the self-attention mechanism is introduced to enhance the classification ability of the system. The experimental results show that in the UCI-HAR data set, the average F_1 score of this network is 98.6%, and that the average accuracy is 98.4%, which is faster than the convolutional neural network-long short-term memory (CNN-LSTM) convergence speed with the accuracy increased by 4.2%, and having a broader application prospect in human activity recognition.

Key words: human activity recognition; sensor; CNN-BiLSTM; self attention

收稿日期: 2023-02-11; 网络首发日期: 2023-04-24

网络首发地址: <http://kns.cnki.net/kcms/detail/61.1294.N.20230424.1418.010.html>

基金项目: 国家自然科学基金资助项目(42105143); 江苏省教育厅, 江苏省高等学校基础科学(自然科学)研究资助项目(580221016)

第一作者: 王赛, 男, 硕士生, 研究方向为可穿戴式设备识别技术。E-mail: 2292917119@qq.com

通信作者: 张立新, 男, 硕士, 高级工程师, 研究方向为计算机技术、嵌入式系统。E-mail: 20201249045@nuist.edu.cn

人类活动识别(HAR)是一项挑战,旨在基于设备交互预测用户活动,它以多种方式帮助人们的日常生活。人类活动检测主要有两种方法:视频图像识别和可穿戴传感器识别。通过视频系统、摄像机识别活动行为,不仅需要安装昂贵的摄像头和基础设施,而且还会因为背景、照明、规模环境等使活动识别变得困难。但是可穿戴传感器,特别是加速度计、陀螺仪等可以将运动转换为识别信号,为移动增加了一个新的维度,比基于视频的识别方法环境约束少,还保护了用户的隐私,在生活援助和老年人看护方面起着更加有效的作用^[1]。

以往的人类活动识别主要基于传统的机器学习算法,研究人员主要根据在一段固定时间内的信号特征以及使用不同的网络算法对人类的活动行为进行识别。比如,Guinness 等^[2]结合加速度计的时域特征数据和四种传统机器学习网络算法依次测试了关于人类活动识别的性能,其中包括 K 邻近(K-nearest neighbor, KNN)、决策树(decision tree, DT)、高斯朴素贝叶斯算法(Gaussian naive Bayesian, GNB)、支持向量机(support vector machine, SVM),实验结果表明在四种算法中, SVM 的性能最好。但是传统的机器学习一般只能使用较为简单的函数形式,不具备复杂的模型表达能力。如果使用复杂的函数学习,那么模型容易产生过拟合,导致网络泛化能力下降。近年来,随着深度学习网络的不断发展,已经逐渐应用于识别和分类等领域,为人类活动识别方面做出了突破,提高了人类活动识别的准确率。He 等^[3]使用加速度计采集数据信息,采用 CNN 网络对走路、跑步等不同的人类活动进行识别,但是单独使用 CNN 网络模型, CNN 中的池化层会丢失大量有价值信息,忽略局部与整体之间关联性。Nguyen 等^[4]使用循环神经网络中的长短期记忆网络(long short-term memory, LSTM)提取活动行为长时间

序列的特征来提升识别的效果;Goh 等^[5]将 CNN 和 LSTM 相结合,建立了可以随时空变换的预测网络模型,提高了活动行为识别的精度,但是 LSTM 按照从前向后的时间序列学习数据内部特征,通过分类器得到分类结果,这样可能会出现一些误判的情况。孙彦玺等^[6]将 CNN 和循环神经网络(recurrent neural network, RNN)的变体 BiLSTM 相结合,对 13 种人类活动行为进行识别,准确率达到 97.7%。

为了进一步提升人类活动识别的准确率,加快网络模型的收敛速度,本文提出了一种基于双通道机制的 CNN-BiLSTM-SA 的混合网络。本网络将加速度和角速度两种不同的传感器数据分别作为输入,有效地保存了各自数据的特征;为了保留数据局部与整体之间关联性,本网络把一维卷积神经网络中的池化层剔除,加入了批量归一化层;然后使用 BiLSTM 网络代替 LSTM 网络,采用双向的时间序列学习特征;最后引入了自注意力机制,增强数据不同部分之间的相关性,进一步提升了系统的分类能力。通过消融实验,证明了本算法比 CNN-LSTM 算法准确率更高,收敛速度更快,并在公开数据集 UCI-HAR 上进行测试,与其他的模型算法进行比较,进一步验证了本网络模型的识别性能。

1 人类活动识别方案

神经网络进行人类活动识别主要分为数据采集、数据预处理、训练模型和验证模型四个步骤,见图 1。基于双通道机制的 CNN-BiLSTM-SA 的网络模型是将加速度计和陀螺仪这两种传感器数据作为输入,分别输入到 CNN 组件层中,经过三层 CNN 组件层的空间特征提取,输出至特征级融合层进行融合,输出至 LSTM 组件层,使用三层 LSTM 组件层提取其时间特征,最后输出至全连接层使用 Softmax 函数激活功能,获得活动分类结果。总体框架见图 2。

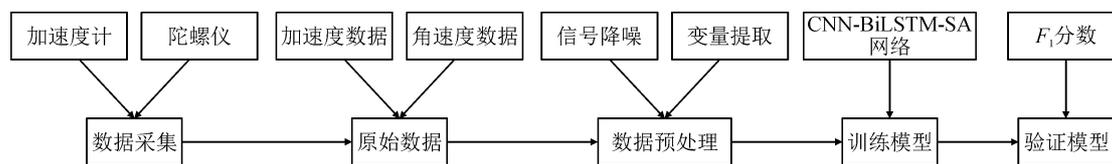


图 1 人类活动识别总体流程图

Fig. 1 Overall flow chart of human activity identification

CNN 组件层由卷积层、批量标准化层和 Dropout 层三部分组成,用于提取数据的空间特征,见图 3。由于第三个 CNN 组件层需要后接特征级融合层,将提取的空间特征进行融合,因此只由一维卷积

层和批量归一化层两层组成。LSTM 组件层是由双向长短期记忆网络层(bi-directional LSTM, BiLSTM),自注意力机制(self-attention, SA)和层标准化层(layer normalization, LN)组成,见图 4。

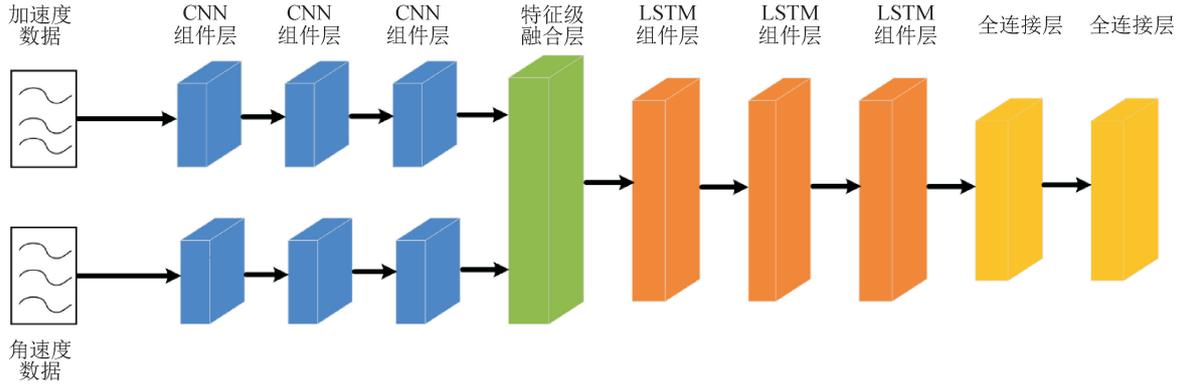


图 2 双通道 CNN-BiLSTM-SA 网络

Fig. 2 Dual-channel CNN-BiLSTM-SA network

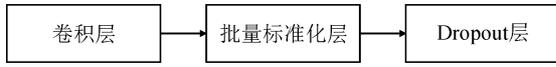


图 3 CNN 组件层

Fig. 3 CNN component layer



图 4 LSTM 组件层

Fig. 4 LSTM component layer

RNN 可以根据深度学习中的先前知识估计当前时间输出,然而由于分解梯度问题,RNN 系统只能在短时间内识别数据^[7]。如果在使用深度学习反向传播策略时不允许梯度深度流动,梯度将被淹没。而 BiLSTM 的细胞状态可用于保存 t 时刻的重要信息,从而保证消息的有效传递,并且相比较 LSTM 的单向输出,其输出的特征是由前向层和反向层共同决定最后输出结果,降低了误判的情况^[8],见图 5。

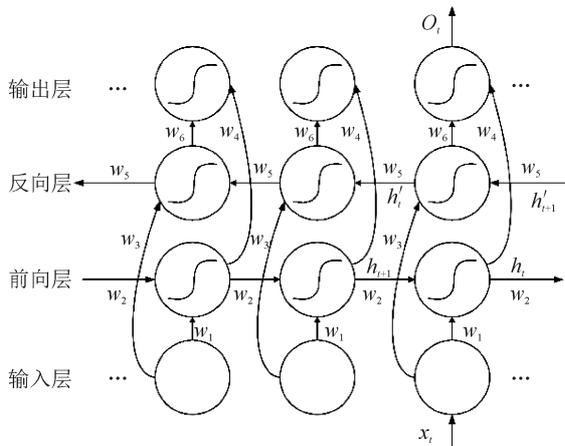


图 5 BiLSTM 网络结构

Fig. 5 BiLSTM network structure

神经元各部分在 t 时刻的运算为:

$$h_t = f(w_1 x_t + w_2 h_{t-1}) \quad (1)$$

$$h'_t = f(w_3 x_t + w_5 h'_{t-1}) \quad (2)$$

$$o_t = g(w_4 x_t + w_6 h'_t) \quad (3)$$

式中: x_t 为 t 时刻的输入; w_i 表示权值矩阵; h_{t-1} 为前一时刻的输出; h'_{t+1} 为下一时刻的输出; h_t 为 t 时刻在前向层的输出; h'_t 为 t 时刻在反向层的输出; O_t 为 t 时刻的输出。

LSTM 组件层中又引入了自注意力机制^[9]。利用三个不同的卷积对输入特征序列进行处理,得到对应的 q, k 和 v 的值。其中 q 和 k 的卷积操作采用相同的参数设置,但是 q 在卷积后利用置换函数进行维度变换。随后将得到的 q 和 k 相乘后经过 sigmoid 激活函数得到权重分数 $\hat{\alpha}$ 。最后,将 $\hat{\alpha}$ 与 v 进行矩阵相乘得到 b , 并将其与输入累加得到最终的输出,见图 6。自注意力机制与 BiLSTM 相结合,增强数据不同部分之间的相关性,进一步提升了系统的分类能力。层标准化层则起到加快网络模型收敛速度的作用,增强了网络模型的泛化性能。

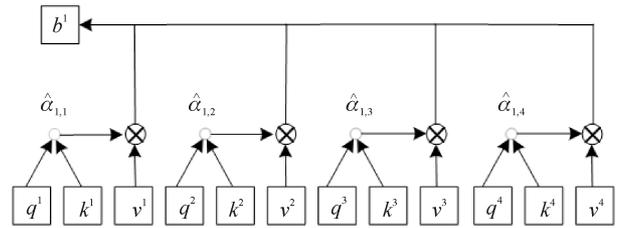


图 6 自注意力计算过程

Fig. 6 Self-attention calculation process

2 实验结果与分析

2.1 数据集采集及其预处理

为了评估提出的网络,采用惯性传感器 MPU6050 采集 X, Y, Z 三轴加速度和三轴角速度数据,通过 I2C 通信传递给 STM32G03 主控芯片,通过 ESP8266 WiFi 模块将数据传送至服务器。服务器则基于这些数据通过网络算法来识别人类活动类型,见图 7。

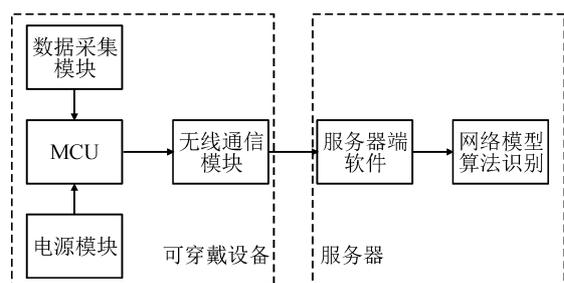


图 7 系统总体框架图

Fig. 7 Overall framework of the system

在日常生活中,由于人体的骨骼结构比较灵活,行人的活动类型多种多样,其重心位置也会随之发生变动,需要寻找最优的人体佩戴位置。相对而言人体腰部的摆动幅度相对更小、更为稳定,因此将可穿戴设备佩戴在人体腰部,采集数据产生的干扰更少,更能客观真实地反映模型算法的性能。因此本实验将传感器佩戴在人体腹部,分别采集了站立,静坐,躺,走,上下楼梯等姿态数据,见图 8。其采样的活动数据点数量见图 9。然后使用了 2015 年更新后的 UCI-HAR 公开数据集作为对比。更新后的 UCI-HAR 数据集包含了从绑在 30 名 19~48 岁年龄段的志愿者腰部的智能手机的加速度计和陀螺仪中提取的 3D(x, y, z)原始信号,共为 12 种日常生活中的行为。

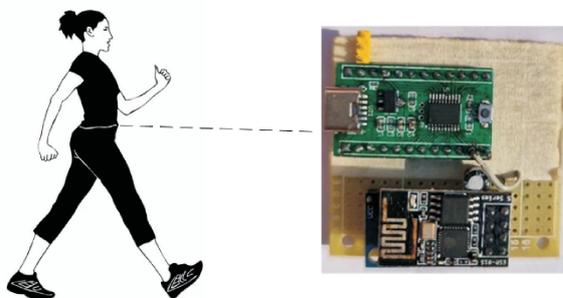


图 8 传感器示意图

Fig. 8 Schematic diagram of sensor

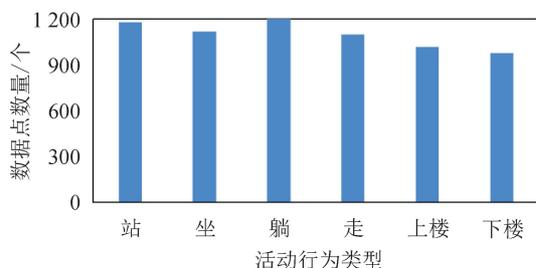


图 9 活动数据点数量

Fig. 9 Number of active data points

传感器信号(加速度计和陀螺仪)中的空缺值和重复值选择直接删除,使用噪声滤波器进行降噪处理,接着滑动窗口设置为 2.56 秒固定宽度进行采

样,每个窗口的重叠设置为 50%,即为 128 个读数/窗口。传感器的加速度信号由重力和人体运动分量组成,通过使用巴特沃斯低通滤波器(Butterworth filter)^[10]将其分离为人体加速度和重力两个部分,滤波器的截止频率设置为 0.3 Hz。提取时频域信号中的标准差,最大值,最小值,加权平均值等特征值。

2.2 超参数选择

本文中的实验算法是使用带有 Tensorflow 后端的 Keras 框架来实现的。网络模型采用的是交叉熵损失函数,可以计算预测值和真实值之间的误差进而评估概率分布之间所存在的差异。学习率使用回调函数的方法自适应设置,其初始学习率为 0.01,当下降停滞时,学习率将达到最佳效果。Dropout 层中丢包率设置为 0.5。使用 Adam 优化器进行迭代轮数为 200,批处理次数为 100 的训练,再使用 AdaGrad 优化器进行迭代轮数为 100 的训练。

网络中两个输入层参数为[(None, 128, 3)],经过三层 CNN 组件层后输出参数为[(None, 128, 48)],再通过特征级融合层后输出参数为[(None, 128, 96)],然后通过三层 LSTM 组合层后输出参数为(None, 24),最后经过两层全连接层后输出参数为(None, 12)。

2.3 实验设置及评估措施

本文所有运行的实验均使用为 Windows 系统,其 CPU 为 AMD Ryzen7 5800H, GPU 为 NVIDIA GeForce RTX 3060 6G 显存。实验使用 Tensorflow 作为一个 Python 库,用于进行神经网络训练、张量运算和自动微分的参数推断,NumPy 用于其他数组操作,Matplotlib 和 Seaborn 用于数据可视化。为验证本实验数据训练的普适性,将 70% 的数据集作为训练集,15% 的数据集作为测试集,15% 的数据集作为验证集。

因为 UCI-HAR 数据集中的活动数据分布不规范,如果分类器将每个实例作为多数类进行预测,并使用总体分类精度(A_{cc})来评估网络输出,则结果将达到高精度,因此总体分类精度不是网络性能评估的良好指标。 F_1 分数(F_1 -score)考虑了假阳性和假阴性,它结合了两个指标“精确度(P_{re})”和“召回率(R_{cc})”,能够正确识别样本总数^[11]。具体公式为:

$$A_{cc} = (T_P + T_N) / (T_P + T_N + F_P + F_N) \quad (4)$$

$$P_{re} = T_P / (T_P + F_P) \quad (5)$$

$$R_{cc} = T_P / (T_P + F_N) \quad (6)$$

$$F_1 = 2P_{re}R_{cc} / (P_{re} + R_{cc}) \quad (7)$$

式中: T_N 、 F_N 分别为真阴性和假阴性的数量; T_P 、 F_P 分别表示真阳性和假阳性的数量。

2.4 消融实验

本文将 CNN-BiLSTM-SA 和 CNN-LSTM 这两种算法模型分别使用自采集数据集中的训练集和 UCI-HAR 公开数据集中的训练集进行实验对比,见图 10 和图 11。

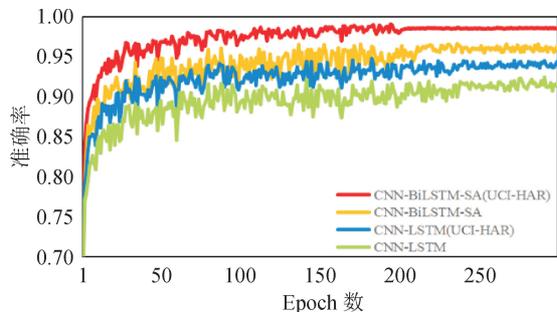


图 10 准确率对比

Fig. 10 Comparison of accuracy

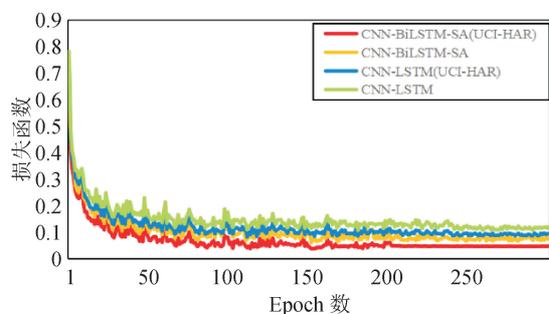


图 11 损失函数对比

Fig. 11 Comparison of loss functions

在 UCI-HAR 数据集和自采集数据集中,这两个网络模型的准确率都是在 1~30 个 Epoch(s)后快速上升,在 30~200 后上升速度变慢,在 200 个 Epochs 后趋于稳定。稳定后在 UCI-HAR 数据集中, CNN-BiLSTM-SA 网络模型的平均准确率大约达到

了 0.984,而 CNN-LSTM 网络模型的平均准确率大约达到 0.942。在自采集数据集中, CNN-BiLSTM-SA 网络模型的平均准确率大约达到了 0.953,而 CNN-LSTM 网络模型的平均准确率大约达到 0.914。因此在准确率识别方面, CNN-BiLSTM-SA 网络模型的性能优于 CNN-LSTM 网络模型。

在 UCI-HAR 数据集和自采集数据集中,这两个网络模型的损失函数都是在 1~30 个 Epoch(s)后快速下降,在 30~200 后下降速度变慢,在 200 个 Epochs 后这两个网络模型基本趋于稳定。稳定后在 UCI-HAR 数据集中, CNN-BiLSTM-SA 网络模型的损失函数大约达到了 0.032,而 CNN-LSTM 网络模型的损失函数大约达到 0.095。在自采集数据集中, CNN-BiLSTM-SA 网络模型的损失函数大约达到了 0.078,而 CNN-LSTM 网络模型的损失函数大约达到 0.116。因此在损失函数方面, CNN-BiLSTM-SA 网络模型的性能优于 CNN-LSTM 网络模型。

使用 CNN-BiLSTM-SA 算法模型通过 UCI-HAR 公开数据集中的验证集和测试集得到的混淆矩阵见图 12 和图 13。混淆矩阵共 12 行 12 列,行标签代表着算法模型预测出的 12 种人类活动行为,列标签代表着算法人类真实活动的 12 种行为。由验证集的混淆矩阵图可知,这 12 中日常活动行为的整体识别率较高,走和下楼这两种活动行为识别率达到了百分百识别。由测试集的混淆矩阵可知,对下楼这种活动行为的识别率达到了百分百识别。但是如图 14 和图 15 所示,因为站和坐的动作摆动幅度很相似,导致了其采样数据的相似度也很高,所以在验证集和测试集的混淆矩阵中,这两种活动行为的混淆程度比较高,识别率比其他活动行为略低一些。

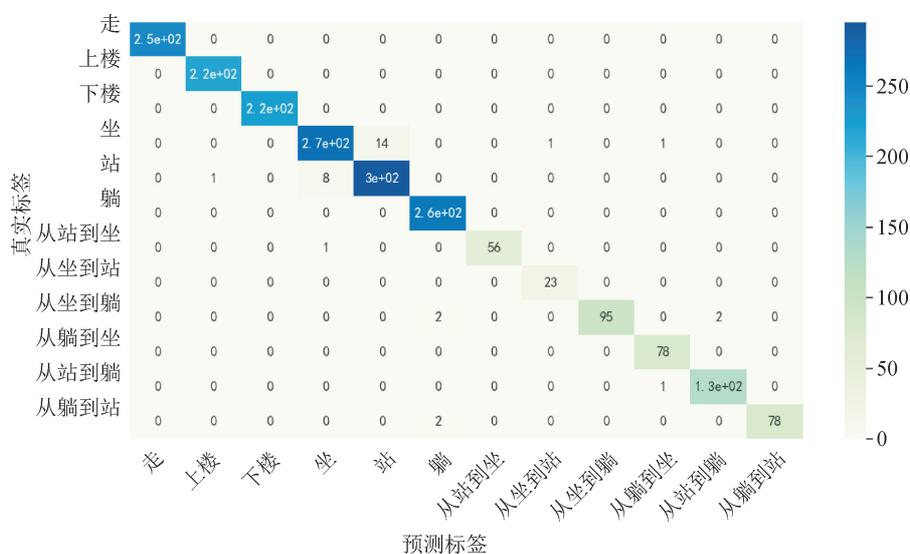


图 12 验证集混淆矩阵

Fig. 12 Confusion matrix of verification set

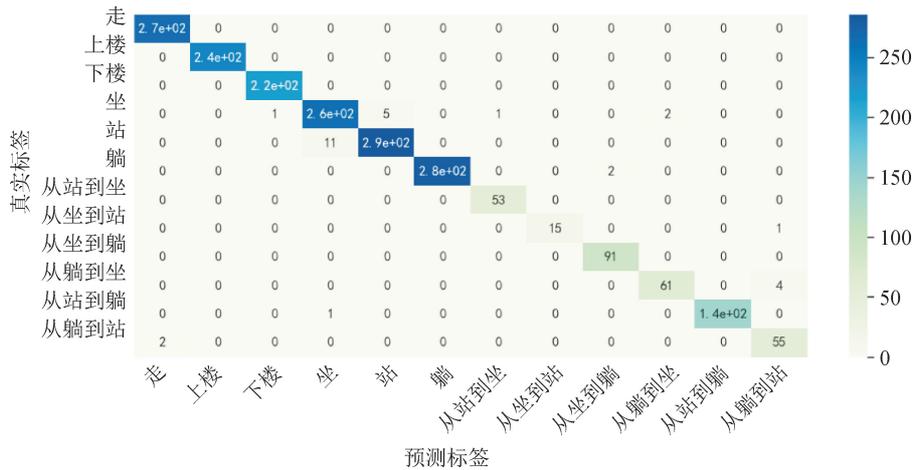


图 13 测试集混淆矩阵

Fig. 13 Confusion matrix of test set

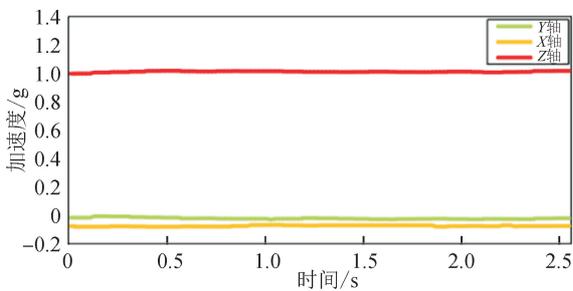


图 14 坐三轴加速度

Fig. 14 Acceleration of sitting three axis

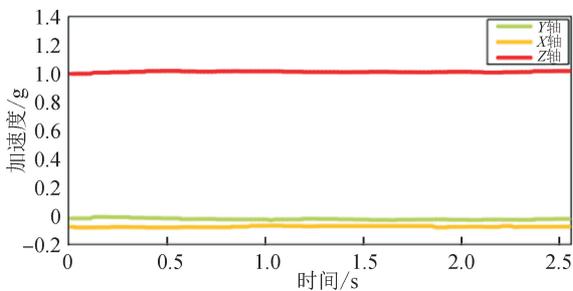


图 15 站三轴加速度

Fig. 15 Acceleration of standing three axis

F_1 分数作为综合评价指标可以客观真实反映出人类活动识别的效果。在 UCI-HAR 公开数据集的测试集和自采集的数据集的测试集中使用 CNN-BiLSTM-SA 算法模型的 F_1 分数见表 1。UCI 数据集的 F_1 分数平均值为 98.6%，其中走、上楼和下楼这三种活动行为的 F_1 分数为 100%。自采集数据集的 F_1 分数平均值为 96.7，其中走这种活动行为的 F_1 分数为 100%。在这两种数据集中，除了坐和站这两种活动行为的 F_1 分数较低以外，其他的活动类型都反映出了高准确率的识别效果。

2.5 不同算法模型的性能对比

在公开数据集 UCI-HAR 的环境中，使用 CNN-BiLSTM-SA 网络模型与其他文章所使用的网络模型在准确率方面进行比较，进一步验证了本算法模型的性能。见表 2，在人类活动识别方面，深度学习比机器学习的识别性能更好，在单一深度学习中，CNN 的性能最好，准确率可以达到 95.2%，而在混合网络模型中，本文中所提出的网络模型，准确率可以达到 98.4%，高于 CNN-LSTM 的 94.1%和 CNN-BiLSTM 网络模型的 97.7%。

表 1 F_1 分数
Tab. 1 F_1 -score

活动类型	F_1 分数/% (UCI 数据集)	F_1 分数/% (自采集数据集)
走	100	100
上楼	100	98
下楼	100	98
坐	96	93
站	96	94
躺	99	97
从站到坐	99	—
从坐到站	98	—
从坐到躺	98	—
从躺到坐	99	—
从站到躺	99	—
从躺到站	99	—
平均值	98.6	96.7

表 2 不同算法模型的准确率对比

算法模型	准确率/%
SVM ^[12]	90.5
LSTM ^[13]	92.8
Stacked LSTM ^[14]	93.13
CNN ^[15]	95.2
CNN-LSTM	94.2
CNN-BiLSTM-SA	98.4

3 结 论

本文提出了双通道 CNN-BiLSTM-SA 深度学习混合网络模型,用于人类日常活动的分类识别。先采用双通道的方式对输入的不同传感器数据提取特征,让每个传感器的数据能保留其独有的特征,再使用 CNN 叠加 BiLSTM 的模式搭建算法模型,有效地提取了数据的深度特征,然后加入自注意力机制,进一步提高了活动识别的效率和准确性,最后使用 UCI-HAR 公开数据集和自采样的数据集,对所提出的算法模型进行验证。实验结果表明,CNN-BiLSTM-SA 算法模型比 CNN-LSTM 算法模型识别效果优秀,收敛速度更快,在 UCI-HAR 数据集中的平均识别准确率达到 98.4%, F_1 分数达到 98.6%,优于其他的算法模型。但是通过混淆矩阵得知本算法模型对坐和站两种活动类型的识别效果仍需改进。

参考文献:

- [1] 何坚,郭泽龙,刘乐园,等. 基于滑动窗口和卷积神经网络的可穿戴人体活动识别技术[J]. 电子与信息学报, 2022,44(1):168-177.
HE Jian, GUO Zelong, LIU Yuanyuan, et al. Human activity recognition technology based on sliding window and convolutional neural network[J]. Journal of Electronics and Information Journal of Electronics & Information Technology, 2022,44 (1): 168-177.
- [2] GUINNESS R E. Beyond where to how: a machine learning approach for sensing mobility contexts using smartphone sensors[J]. Sensors, 2015, 15: 9962-9985.
- [3] HE Jian, ZHANG Zihao, WANG Xiaoyi, et al. A low power fall sensing technology based on FD-CNN[J]. IEEE Sensors Journal, 2019, 19(13): 5110-5118.
- [4] NGUYEN T D T, PARK J H, HOSSAIN M I, et al. Performance analysis of data parallelism technique in

machine learning for human activity recognition using LSTM[C]//2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), IEEE, 2019: 387-391.

- [5] GOH J X, LIM K M, LEE C P. 1D convolutional neural network with long short-term memory for human activity recognition[C]//2021 IEEE International Conference on Artificial Intelligence in Engineering and Technology (ICAET), IEEE, 2021: 1-6.
- [6] 孙彦玺,陈继斌,武东辉. 基于卷积神经网络-双向长短期记忆网络的人体活动识别方法[J]. 科学技术与工程, 2022,22(4):1517-1525.
SUN Yanxi, CHEN Jibin, WU Donghui. Human activity recognition method based on convolutional neural network-bidirectional long short term memory network [J]. Science Technology and Engineering, 2022, 22 (4): 1517-1525.
- [7] 杨万鹏,李擎,雷明. 人体动作识别的特征级融合 LSTM-CNN 方法研究[J]. 电子测量技术, 2021, 44 (17):173-180.
YANG Wanpeng, LI Qing, LEI Ming. Research on feature-level fusion LSTM-CNN method for human activity recognition [J]. Electronic Measurement Technology, 2021,44 (17): 173-180.
- [8] LI Yong, WANG Luping. Human activity recognition based on residual network and BiLSTM[J]. Sensors, 2022, 22(2): 635.
- [9] 王金甲,周雅倩,郝智. 基于注意力模型的多传感器人类活动识别[J]. 计量学报, 2019,40(6):958-969.
WANG Jinjia, ZHOU Yaqian, HAO Zhi. Multi-sensor human activity recognition based on attention model [J]. Acta Metrologica Sinica, 2019, 40(6): 958-969.
- [10] 郝占军,张岱阳,党小超,等. 基于信道状态信息的非接触式人员动作识别方法[J]. 计算机工程, 2021,47(6): 172-181.
HAO Zhanjun, ZHANG Daiyang, DANG Xiaochao, et al. Non-contact human motion recognition method based on channel state information [J]. Computer Engineering, 2021,47 (6): 172-181.
- [11] 张琳,易卿武,黄璐,等. 基于 Wavelet-CNN 网络的人类活动识别技术 [J]. 无线电工程, 2022, 52 (4): 590-597.
ZHANG Lin, YI Qingwu, HUANG Lu, et al. Human activity recognition based on wavelet-CNN architecture [J]. Radio Engineering, 2022, 52 (4): 590-597.

(下转第 290 页)

- on the emerging classical vision [J]. *On Economic Problems*, 2020(8): 10-19.
- [28] 唐晓华,张欣珏,李阳. 中国制造业与生产性服务业动态协调发展实证研究[J]. *经济研究*, 2018, 53(3): 79-93.
TANG Xiaohua, ZHANG Xinyu, LI Yang. Dynamic coordination development in China's manufacturing and manufacturing-related service industries[J]. *Economic Research Journal*, 2018, 53(3): 79-93.
- [29] 夏飞龙,朱丽萍. 财政分权、政商关系和结构性产能过剩——基于2001-2011年中国工业面板数据和GMM的分析[J]. *商业研究*, 2022(1): 76-84.
XIA Feilong, ZHU Liping. Fiscal decentralization, government-business relationship and structural overcapacity: based on industrial panel data in China from 2001 to 2011 and GMM[J]. *Commercial Research*, 2022(1): 76-84.
- [30] 李健,冯会迎. 高技术制造业与高技术服务业协同集聚的经济增长效应[J]. *科技进步与对策*, 2020, 37(17): 54-62.
LI Jian, FENG Huiying. The economic growth effect of the high-tech manufacturing industry and high-tech service industry co-agglomeration[J]. *Science & Technology Progress and Policy*, 2020, 37(17): 54-62.
- [31] 赵星,赵仁康,董帮应. 国际产业转移与贸易保护损害的关系研究[J]. *经济问题探索*, 2013(11): 130-138.
ZHAO Xing, ZHAO Renkang, DONG Bangying. The research on the relationship of international industrial transfer and damaging trade protection [J]. *Inquiry into Economic Issues*, 2013(11): 130-138.
- [32] HOLMES T J, STEVENS J J. Geographic concentration and establishment scale[J]. *The Review of Economics and Statistics*, 2002, 84(4): 682-690.
- [33] 阎川,雷婕. 财政分权对产业集聚影响的实证分析[J]. *经济评论*, 2019(3): 104-122.
YAN Chuan, LEI Jie. An empirical analysis of the influence of fiscal decentralization on industrial agglomeration[J]. *Economic Review*, 2019(3): 104-122.
- [34] 王旭娜,谭清美. 生产性服务嵌入对制造业竞争力提升的实证研究——基于中国区域城市的空间计量分析[J]. *科技管理研究*, 2021, 41(5): 135-143.
WANG Xu'na, TAN Qingmei. An empirical study on the improvement of manufacturing competitiveness by producer service embedding: the spatial econometric analysis based on Chinese regional cities[J]. *Science and Technology Management Research*, 2021, 41(5): 135-143.
- [35] ALFARO L, CHANDA A, SEBNEK K O, et al. FDI and economic growth: the role of local financial markets [J]. *Journal of International Economics*, 2004, 64(1): 89-112.
- [36] 黄永明,陈宏. 外商直接投资对中国产业升级影响的研究[J]. *管理现代化*, 2018, 38(4): 22-25.
HUANG Yongming, CHEN Hong. Study on FDI affecting the upgrading of Chinese industry[J]. *Modernization of Management*, 2018, 38(4): 22-25.
- [37] 杨玲. 生产性服务贸易进口复杂度对上海科技企业技术水平的异质性影响效用研究[J]. *中央财经大学学报*, 2014(9): 107-112.
YANG Ling. Producer service import complexity improve shanghai firms' technology progress[J]. *Journal of Central University of Finance & Economics*, 2014 (9): 107-112.

(责任编辑 周 蓓)

(上接第259页)

- [12] DESHPNANDE A, WARHADE K K. An improved model for human activity recognition by integrated feature approach and optimized SVM[C]//2021 International Conference on Emerging Smart Computing and Informatics (ESCI), IEEE, 2021: 571-576.
- [13] ADI S E, CASSON A J. Design and optimization of a Tensor Flow Lite deep learning neural network for human activity recognition on a smartphone[C]//2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2021: 7028-7031.
- [14] ULLAH M, ULLAH H, KHAN S D, et al. Stacked LSTM network for human activity recognition using smartphone data[C]//2019 8th European Workshop on Visual Information Processing (EUVIP), IEEE, 2019: 175-180.
- [15] JUNAGADE N, KULKARNI S. Human activity identification using CNN[C]//2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), IEEE, 2020: 1058-1062.
(责任编辑 王绪迪)