

DOI:10.19651/j.cnki.cmt.2107170

基于 CNN-Mogrifier LSTM 的人体运动模式识别算法

李浩^{1,2} 于志远^{1,2} 尹业成^{1,2} 闫国栋^{1,2}

(1.北京精密机电控制设备研究所 北京 100076; 2.航天伺服驱动与传动技术实验室 北京 100076)

摘要: 随着传感器、微电子等技术的发展,通过可穿戴式传感器对人体的运动模式进行识别,具有广泛的应用价值,如何提高识别的准确率,具有重要研究意义。考虑到人体下肢运动的特点,提出了一种基于 CNN 和 Mogrifier LSTM 的人体运动模式识别算法,先利用 CNN 提取原始数据的局部相关特征,再使用 Mogrifier LSTM 代替全连接层,挖掘局部相关特征的前后依赖关系,对行走、跑步、上楼梯、下楼梯、上坡和下坡 6 种常见的运动模式进行识别。实验结果表明,相比于传统 LSTM 算法,Mogrifier LSTM 的准确率提升了 1.03%,将 CNN 和 Mogrifier LSTM 相结合后,准确率进一步提升了 1.17%,达到了 98.18%,证明了算法的优越性。

关键词: 形变长短时记忆网络;卷积神经网络;人体运动模式识别;惯性测量单元

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Human motion pattern recognition algorithm based on CNN-Mogrifier LSTM

Li Hao^{1,2} Yu Zhiyuan^{1,2} Yin Yecheng^{1,2} Yan Guodong^{1,2}

(1. Beijing Research Institute Precision Mechatronics and Controls, Beijing 100076, China;

2. Laboratory of Aerospace Servo Actuation and Transmission, Beijing 100076, China)

Abstract: With the development of sensor technology and microelectronics, it has a wide application value to recognize human motion patterns by wearable sensors. It is of great significance to improve the accuracy of recognition. This paper considered the characteristics of human lower limb motion, and proposed a human motion pattern recognition algorithm based on CNN and Mogrifier LSTM. First, CNN is used to extract local related features of the original data, then Mogrifier LSTM is used to replace the full connection layer to mine the front and back dependencies of local related features. Recognize the six common motion patterns of walking, running, upstairs, downstairs, uphill and downhill. The experimental results show that compared with the traditional LSTM algorithm, the accuracy of Mogrifier LSTM is improved by 1.03%. After combining CNN and Mogrifier LSTM, the accuracy is further improved by 1.17% to 98.18%, which proves the superiority of the algorithm proposed.

Keywords: mogrifier LSTM; convolutional neural networks; human motion patterns recognize; inertial measurement unit

0 引言

通过可穿戴式传感器对人体的运动模式进行识别,主要是指利用可穿戴于人体的传感器采集肢体在运动过程中产生的数据,通过对传感器数据进行分析,进而识别出穿戴者当前的运动模式,如上楼梯、行走、跑步等^[1],在外骨骼机器人^[2]、智能假肢^[3]、智能家居^[4]等领域具有重要的研究意义。

人体运动模式识别通常包括数据采集与预处理、特征提取、训练分类器与识别 4 个阶段。目前常见的思路主要

为基于传统机器学习的方法和基于深度学习的方法两种:基于传统机器学习的方法需要手动从传感器原始数据中选择相关特征,如均值、标准差、组内相关系数、最大值、最小值等^[5],再利用人工选择的特征训练分类器,如支持向量机(support vector machine, SVM)^[6]、线性判别分析(linear discriminant analysis, LDA)^[7]、随机森林(random forest, RF)^[8]等。基于传统机器学习方法的识别性能在很大程度上取决于特征的相关程度,若特征无法体现出各运动模式之间的差异性,识别效果会大幅降低。近年来,随着深度学习技术的不断发展,基于深度学习的方法逐渐成为人体运

动模式识别的主流研究方向^[9],与传统机器学习相比,深度学习可以通过隐藏层网络自动进行特征提取,从而提升识别的速度和精度。Ronao 等^[10]提出采用卷积神经网络(convolutional neural networks, CNN)对行走、上下楼梯等模式进行识别,效果优于 SVM。Wang 等^[11]利用长短时记忆网络(long short-term memory, LSTM)具有记忆和存储功能的特点,进行运动模式的识别。Francisco 等^[12]提出了一种 CNN 和 LSTM 相结合的混合模型,将 CNN 的输出做为 LSTM 的输入。Ahmad 等^[13]利用 CNN-LSTM 混合模型对行走、上楼梯等 6 种运动模式进行识别,识别效果好于 SVM 和 KNN。

虽然目前关于通过可穿戴传感器对人体的运动模式进行识别的研究已经取得了一定的成果,但受到步速、不同穿戴者行为习惯差异等因素的影响,各运动模式间具有显著类间相似性和类内可变性^[14],部分运动模式的区分难度较大,如何进一步提升识别的精度,仍面临许多挑战。

考虑人体下肢常见的运动模式需要双腿协作完成,数据具有明显的前后依赖关系,本文提出了一种将 CNN 和 Mogrifier LSTM 相结合的混合模型,简称 CNN-Mogrifier LSTM 算法。该算法先利用 CNN 自动提取各个传感器数据之间的潜在相关性,生成特征向量,然后将特征向量做为输入,利用 Mogrifier LSTM 充分挖掘特征之间的前后依赖关系,具有较强的识别性能。

1 算法设计

1.1 数据采集与预处理

惯性测量单元(inertial measurement unit, IMU)由加速度计和陀螺仪组成,可以测量下肢在三维空间中的加速度和角速度。为了尽可能获取较多的运动信息,本文将 IMU 分别安装于左右大小腿外侧,安装位置如图 1 所示。

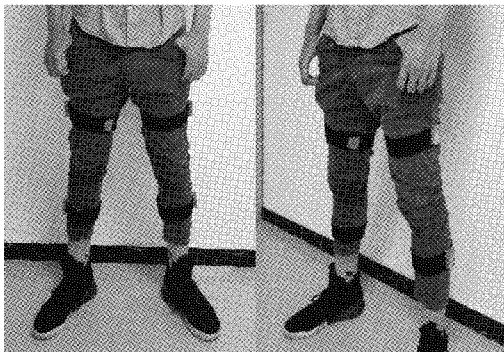


图 1 传感器安装位置

经调查研究发现,人体在执行不同的运动模式时,沿重力方向的加速度与矢状面角速度的差异最为明显,因此本文采集 4 个 IMU 沿重力方向的加速度和矢状面角速度,共 8 轴数据。IMU 采集到的数据时间跨度较大,无法直接用于识别,需要进行数据分割。本文选用滑动窗口法提取用于识别的数据样本,如图 2 所示,首先设定长度为 S 的时间

窗口,然后将窗口按固定步长 L 沿数据进行滑动,每一个滑动窗口内的数据均代表一个识别样本, L 取值通常为 S 的 $1/2$ 。

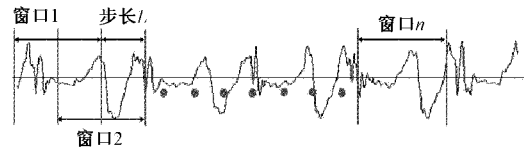


图 2 滑动窗口法

一方面,在信号采集过程中,受传感器生产工艺、肢体不自主的抖动等因素的影响,难以采集到纯净的惯性信号;另一方面,加速度和角速度的量纲差异较大,不能将其直接混合做为模型的输入,为了消除噪声和量纲差异的影响,需要将窗口中传感器原始数据进行滤波和归一化操作。本文选择基于最小二乘原理的 Savitzky-Golay 滤波方式,数据滤波前后的效果如图 3 所示,可以看出,Savitzky-Golay 滤波可以较好地消除噪声的影响。

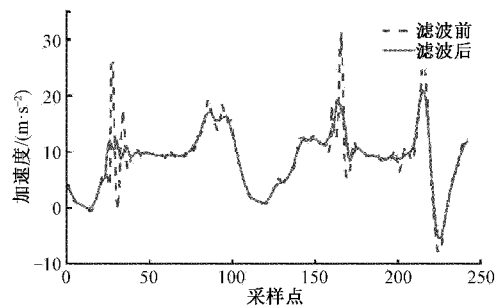


图 3 Savitzky-Golay 滤波

1.2 卷积神经网络

卷积神经网络是一种前馈神经网络,具有局部连接特性和权值共享特性,能够自动对原始数据进行局部空间特征提取,广泛应用于图像识别等领域^[15]。CNN 的主要组成结构包括卷积层、池化层和全连接层。卷积层利用不同的卷积核对输入数据进行卷积运算,从而实现特征的提取。卷积层计算公式如下:

$$y_n = f\left(\sum_{i=1}^M [(W_n^i \cdot x_i) + b_n]\right) \quad (1)$$

其中, x_i 为当前卷积层的输入数据, b_n 为偏置项, W_n 为权重矩阵, f 为激活函数, y_n 为当前卷积层提取到的第 n 个特征。卷积层输出的参数数量往往较大,不便于训练,池化层主要用于数据压缩,在减小计算量的同时降低过拟合风险,通常包括最大池化和平均池化两种。最后将原始数据经过卷积和池化操作生成的特征向量送入全连接层进行处理,便可得到 CNN 的识别结果。CNN 模型结构如图 4 所示。

1.3 Mogrifier LSTM

CNN 虽然可以获取数据的局部相关特征,但难以挖掘数据的前后依赖关系。LSTM 是循环神经网络(recurrent

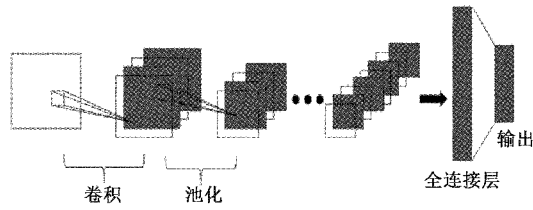


图 4 CNN 结构

neural network, RNN)的一种衍生网络,通过引入“记忆单元”和“门”机制,解决了传统 RNN 在处理较长时间序列数据中存在的梯度爆炸和梯度消失的问题^[16],能够充分挖掘长期时间序列数据的前后依赖关系,LSTM 的结构如图 5 所示。

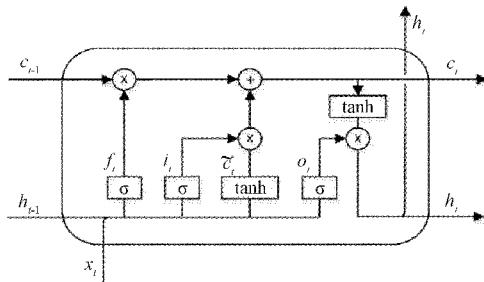


图 5 LSTM 结构

LSTM 中的“门”机制包括输入门 i_t 、遗忘门 f_t 和输出门 o_t , x_t 为当前节点的输入, h_{t-1} 为前一节点的输出, c_{t-1} 为前一节点的状态。LSTM 通过输入门 i_t 对当前节点的输入 x_t 和 h_{t-1} 进行选择记忆,决定有多少信息可以存储到当前节点以生成新状态 c_t ,然后通过遗忘门 f_t 对前一节点的状态 c_{t-1} 进行选择遗忘,只保留部分有用信息,最后通过输出门 o_t 将 c_t 转换为当前节点的输出 h_t 。LSTM 的具体更新过程如下:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (3)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (4)$$

$$\tilde{c}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (5)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (6)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (7)$$

其中, W 为各个门的权重系数矩阵, b 为偏置矩阵, σ 为 sigmoid 激活函数。由式(7)可以看出,当前节点输出的状态 c_t 由前一节点的状态和当前节点的内部信息加权生成,只要遗忘门 f_t 不为 0, LSTM 就可以对之前节点的信息进行记忆。

由图 5 可以观察到,当前节点的输入 x_t 和前一节点的输出 h_{t-1} 在送入节点前并无直接联系。Melis 等^[17]认为相互独立的 x_t 和 h_{t-1} 会损失数据的部分时间相关性,因此在 LSTM 的基础上提出了 Mogrifier LSTM。Mogrifier LSTM 并不改变 LSTM 的原有结构,而是先将 x_t 和 h_{t-1} 先按式(8)进行 r 轮充分交互,以增强时间序列中的主要特

征、减弱次要特征,然后再将更新后的 x_t 和 h_{t-1} 送入当前节点,进而提升模型提取数据时间相关性的能力,图 6 为 $r=5$ 时的 Mogrifier LSTM 示意图。其中 x^{-1} 为当前节点的输入 x_t , h^0 为上一节点的输出 h_{t-1} ,超参数 r 为 x_t 和 h_{t-1} 交互的轮数,当 $r=0$ 时, Mogrifier LSTM 退化为 LSTM, $i \in [1, \dots, r]$, 当 i 为奇数时,对 x_t 进行更新,当 i 为偶数时,对 h_{t-1} 进行更新, Q_i 和 R_i 为模型的参数矩阵。文献[17]通过实验,证明了在文本任务领域, Mogrifier LSTM 的性能明显优于 LSTM。

$$x^i = 2\sigma(Q^i h_{t-1}^{i-1})x^{i-2}, \text{ for odd } i \in [1, \dots, r] \quad (8)$$

$$h_{t-1}^i = 2\sigma(R^i x^{i-1})h_{t-1}^{i-2}, \text{ for even } i \in [1, \dots, r]$$

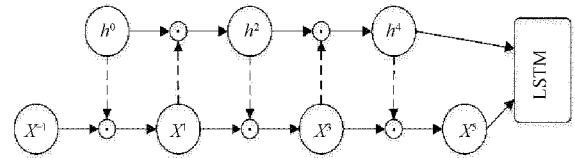


图 6 Mogrifier LSTM

1.4 CNN-Mogrifier LSTM 混合模型

为了从传感器数据中挖掘出更多的信息,提升识别精度,本文将 Mogrifier LSTM 用于人体运动模式识别领域,并与 CNN 相结合,提出了一种基于 CNN-Mogrifier LSTM 的识别算法,结构如图 7 所示。

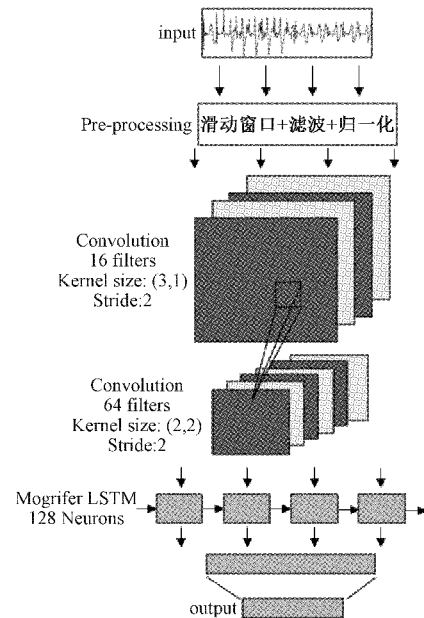


图 7 CNN-Mogrifier LSTM 算法结构

CNN-Mogrifier LSTM 算法首先利用滑动窗口法对各 IMU 采集到的数据进行划分,并对窗口中的原始数据进行滤波、归一化操作,然后通过两个卷积层提取数据的局部空间特征,再将卷积层提取的特征做为 Mogrifier LSTM 层的输入,通过对 x_t 和 h_{t-1} 进行充分交互,充分挖掘局部空间特征前后依赖关系,最后利用全连接层和 Softmax 函数对

当前数据所属的运动模式进行识别。为了包含识别运动模式所需要的信息,窗口长度需要在 200 ms 以上^[18],因此本文滑动窗口参数 S 选择 30 (300 ms), L 选择 15 (150 ms),模型的 $\text{batch size} = 128$, $\text{Dropout} = 0.3$,结构参数如表 1 所示。

表 1 模型结构参数

编号	层类型	输出维度	核大小	步长
1	Input layer	30×8	—	—
2	1 st Convolution layer	$16 \times 30 \times 8$	3,1	2
3	2 st Convolution layer	$64 \times 16 \times 5$	2,2	2
4	Mogrifier LSTM layer	128	—	—
5	Dropout	—	—	—
6	1 st Full-connected layer	32	—	—
7	2 st Full-connected layer	6	—	—

2 实验与结果分析

2.1 实验平台

本文使用的惯性测量单元为荷兰 Xsens 公司的 Xsens MTw。实验所用标准民用楼梯每阶的高度为 16 cm,宽 27 cm,斜坡坡度约为 8° 。电脑处理器为 Intel(R) Core(TM) i7-8550U CPU @ 1.80 GHz 2.00 GHz, Windows10 64 位操作系统,模型基于 Pytorch 框架实现。

2.2 数据集构造

为了验证模型的识别效果,本文选择 10 名测试人员进行实验。测试人员的运动模式依次为行走-上下楼梯-跑步-上下坡,通过滑动窗口法,共采集到 27 040 个实验样本,选择其中 18 319 个样本(约 70%)做为训练集,剩余样本(约 30%)做为测试集进行实验,数据集中各运动模式的样本数如表 2 所示。

表 2 实验数据集

运动模式	样本总数	训练集	测试集
行走	5 151	3 673	1 478
跑步	4 723	3 630	1 093
上楼梯	4 242	2 913	1 329
下楼梯	3 786	2 571	1 215
上坡	4 761	3 346	1 415
下坡	4 377	3 056	1 321

2.3 结果与分析

本文分别采用 LSTM、Mogrifier LSTM、CNN-LSTM 和本文提出的 CNN-Mogrifier LSTM 4 种算法进行实验,选择混淆矩阵做为算法性能的评价标准,混淆矩阵中纵轴标签为样本的真实运动模式,横轴标签为模型识别结果。

Mogrifier LSTM 中 x_t 和 h_{t-1} 的交互轮数 r 需通过实验确定,图 8 为 r 在不同取值下 Mogrifier LSTM 的识别准

确率,可以发现, $r=5$ 时的准确率最高,因此本文 Mogrifier LSTM 和 CNN-Mogrifier LSTM 中交互次数 r 均选择 5,图 9 为各算法得到的混淆矩阵,表 3 为各算法得到的识别准确率。

对图 9 和表 3 进行分析可以发现,将 x_t 和 h_{t-1} 进行 5 轮交互后,Mogrifier LSTM 可以更好挖掘时间序列数据的前后依赖关系,识别准确率明显优于 LSTM。与仅使用

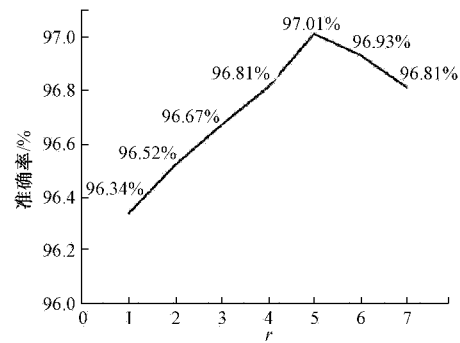
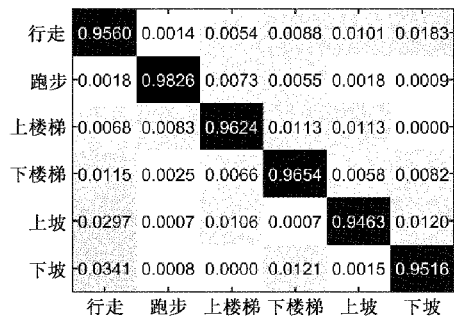
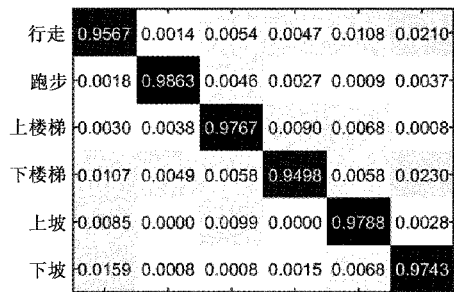


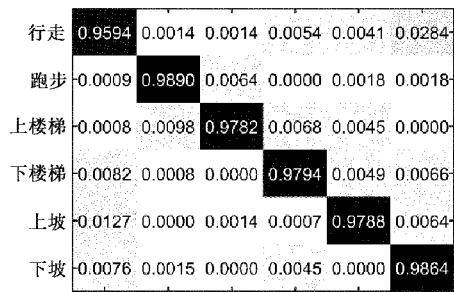
图 8 r 不同取值时 Mogrifier LSTM 识别准确率



(a) LSTM



(b) Mogrifier LSTM



(c) CNN-LSTM

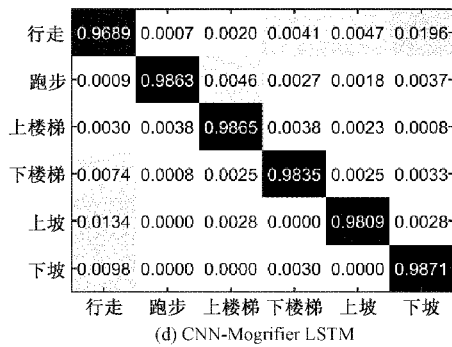


图 9 混淆矩阵

表 3 识别准确率 %

LSTM	Mogrifier LSTM	CNN-LSTM	CNN-Mogrifier LSTM
95.98 (7 535/7 851)	97.01 (7 616/7 851)	97.78 (7 677/7 851)	98.18 (7 708/7 851)

LSTM 相比, CNN-LSTM 模型先通过 CNN 提取数据的大量局部空间特征, 可以为 LSTM 提供更多的有效输入数据, 进而提升识别的准确率。而本文所提出的 CNN-Mogrifier LSTM 模型与 CNN-LSTM 模型相比, 充分利用了 Mogrifier LSTM 的优势, 可以比 LSTM 更好地挖掘局部空间特征的前后依赖关系。实验结果表明, CNN-Mogrifier LSTM 算法取得了较高的识别准确率。

3 结 论

本文提出了一种将 CNN 和 Mogrifier LSTM 相结合的深度学习算法, 用于通过可穿戴式传感器进行人体的运动模式识别。该算法充分利用了 Mogrifier LSTM 的优势, 对 CNN 提取的局部空间特征的前后依赖关系进行深度挖掘。实验结果表明, 本文所提的基于 CNN-Mogrifier LSTM 方法取得了 98.18% 的准确率, 优于 LSTM、Mogrifier LSTM 和 CNN-LSTM, 证明了算法的优越性。但由于实验所用的斜坡较为平缓, 行走、上坡和下坡仍有少量的误识别情况, 未来工作将采用 GRU 或 Bi-LSTM 等 LSTM 变体与 Mogrifier 思想和 CNN 相结合, 设计新的算法, 进一步提升感知精度。

参考文献

- [1] 路永乐, 张欣, 龚爽, 等. 基于 MEMS 惯性传感器的人体多运动模式识别[J]. 中国惯性技术学报, 2016, 24(5): 589-594.
- [2] 高经纬, 马超, 姚杰, 等. 基于机器学习的人体步态检测智能识别算法研究[J]. 电子测量与仪器学报, 2021, 35(3): 49-55.
- [3] 李向新, 田岚, 郑悦, 等. 一种下肢运动意图识别算法性能实时测评系统[J]. 仪器仪表学报, 2020, 41(5):

- 99-107.
- [4] ISHIMARU S, HOSHIKA K, KUNZE K, et al. Towards reading trackers in the wild: detecting reading activities by EOG glasses and deep neural networks [C]. Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, 2017: 704-711.
- [5] 苏本跃, 王婕, 刘双庆, 等. 惯性动捕数据驱动下的智能下肢假肢运动意图识别方法[J]. 自动化学报, 2020, 46(7): 1517-1530.
- [6] 杨智超, 李国辉, 李伟韵, 等. 基于分散熵和支持向量机的运动状态识别[J]. 国外电子测量技术, 2019, 38(7): 28-31.
- [7] 张乾勇, 张涛, 赵治羽. 基于多传感器的人体运动模式识别[J]. 传感器与微系统, 2019, 38(2): 73-76.
- [8] BALLI S, SAGBAS E A, PEKER M. Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm[J]. Measurement and Control, 2019, 52(1/2): 37-45.
- [9] WANG J, CHEN Y, HAO S, et al. Deep learning for sensor-based activity recognition: A survey [J]. Pattern Recognition Letters, 2019, 119: 3-11.
- [10] RONAO C A, CHO S B. Deep convolutional neural networks for human activity recognition with smartphone sensors [C]. International Conference on Neural Information Processing, Springer, Cham, 2015: 46-53.
- [11] WANG L K, LIU R Y. Human activity recognition based on wearable sensor using hierarchical deep LSTM networks [J]. Circuits, Systems, and Signal Processing, 2020, 39(2): 837-856.
- [12] FRANCISCO O, DANIELL R. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition [J]. Sensors, 2016, 16(1): 115.
- [13] AHMAD W, KAZMI B M, ALI H. Human activity recognition using multi-head CNN followed by LSTM [C]. 2019 15th International Conference on Emerging Technologies (ICET), IEEE, 2019: 1-6.
- [14] 李仲, 刘瑛奇, 管小荣. 步速对人体下肢运动特性影响的实验研究[J]. 机械设计与制造, 2019(12): 235-238, 242.
- [15] 张异凡, 黄亦翔, 汪开正, 等. 用于心律失常识别的 LSTM 和 CNN 并行组合模型[J]. 哈尔滨工业大学学报, 2019, 51(10): 76-82.

- [16] HAMMERLA N Y, HALLORAN S, PLOETZ T. Deep, convolutional, and recurrent models for human activity recognition using wearables [J]. Journal of Scientific Computing, 2016, 61(2):454-476.
- [17] MELIS G, KOCISKY T, BLUNSOM P. Mogrifier LSTM [C]. International Conference on Learning Representations, 2019, ArXiv: 1909.01792.
- [18] 宋戈,叶晓东,王美玲.一种基于轻量级梯度提升机的意图识别方法[J].计算机仿真,2019,36(10):312-315,343.

作者简介

李浩,硕士研究生,主要研究方向为智能机构技术。

E-mail: lihao_calt18@163.com

于志远,硕士研究生导师,研究员,主要研究方向为智能机构技术。

E-mail: happytimmyu@163.com

尹业成,工程师,主要研究方向为机器人及机器学习。

E-mail: yycheng_work@126.com

闫国栋,工程师,主要研究方向为深度学习。

E-mail: yanguodongbit@163.com