

基于 CNN-BiGRU 的复杂连续人体活动 Wi-Fi 感知方法

刘洋^{1,2}, 董安明^{1,2,3}, 禹继国^{3,4}, 赵恺⁵, 周酉⁶

1. 齐鲁工业大学(山东省科学院)山东省计算中心(国家超级计算济南中心)算力互联网与信息安全教育部重点实验室, 山东 济南 250353;
2. 齐鲁工业大学(山东省科学院)计算机科学与技术学院, 山东 济南 250353;
3. 齐鲁工业大学(山东省科学院)大数据研究院, 山东 济南 250353;
4. 山东省基础科学研究中心(计算机科学)山东省计算机网络重点实验室, 山东 济南 250353;
5. 中国科学院自动化研究所, 北京 100190;
6. 山东海看新媒体研究院有限公司, 山东 济南 250013)

摘要: 基于 Wi-Fi 信道状态信息(CSI, channel state information)的人体活动感知在虚拟现实、智能游戏、元宇宙等未来智能交互场景具有重要的应用前景, 复杂连续人体活动的精准感知是 Wi-Fi 感知的重要挑战。卷积神经网络(CNN, convolutional neural network)具备空间特征提取能力, 但对数据的时序特征建模能力差。而适用于时间序列数据建模的长短期记忆(LSTM, long short-term memory)网络或门控循环单元(GRU, gated recurrent unit)网络忽视了对数据空间特征的学习。针对此问题, 提出了一种融合双向门控循环单元(BiGRU, bidirectional gated recurrent unit)网络的改进型 CNN。所提网络利用 BiGRU 的双向特征提取能力捕捉时序数据前后信息的关联和依赖性, 实现时序 CSI 数据的时空特征提取, 进而呈现动作与 CSI 数据的映射关系, 从而提高对复杂连续动作的识别精度。以篮球动作为场景对所提网络结构进行了实验, 结果表明, 该方法在多种条件下识别准确率均高于 95%, 与传统多层感知机(MLP, multi-layer perceptron)、CNN、LSTM、GRU、具有注意力机制的双向长短期记忆(ABLSTM, attention based bidirectional long short-term memory)网络等基线方法相比, 识别准确率提升了 1%~20%。

关键词: 信道状态信息; 人体活动感知; 复杂连续活动; 卷积神经网络; 双向门控循环单元

中图分类号: TN92

文献标志码: A

doi: 10.11959/j.issn.2096-3750.2023.00360

A Wi-Fi sensing method for complex continuous human activities based on CNN-BiGRU

LIU Yang^{1,2}, DONG Anming^{1,2,3}, YU Jiguo^{3,4}, ZHAO Kai⁵, ZHOU You⁶

1. Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China
2. School of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China
3. Big Data Research Institute, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China

收稿日期: 2023-02-27; **修回日期:** 2023-07-01

通信作者: 董安明, anmingdong@qlu.edu.cn

基金项目: 国家重点研发计划(No.2019YFB2102600); 国家自然科学基金资助项目(No.61701269, No.62272256); 山东省科技型中小企业创新能力提升工程(No.2022TSGC2180, No.2022TSGC2123); 济南市“高校 20 条”自主培养创新团队项目(No.202228093); 齐鲁工业大学(山东省科学院)科教产融合试点工程项目(基础研究类)先导项目(No.2022XD001); 齐鲁工业大学(山东省科学院)计算机科学与技术学科基础研究加强计划(No.2021JC02014)

Foundation Items: The National Key Research and Development Program (No.2019YFB2102600), The National Natural Science Foundation of China (No.61701269, No.62272256), The Innovation Capability Enhancement Program for Small and Medium-sized Technological Enterprises of Shandong Province (No.2022TSGC2180, No.2022TSGC2123), The Innovation Team Cultivating Program of Jinan (No.202228093), The Pilot Engineering Project of Science, Education and Industry Integration (Basic Research) of Qilu University of Technology (Shandong Academy of Sciences) (No.2022XD001), The Basic Research Strengthening Program of Computer Science and Technology Discipline of Qilu University of Technology (Shandong Academy of Sciences) (No.2021JC02014)

4. Shandong Provincial Key Laboratory of Computer Networks, Shandong Fundamental Research Center for Computer Science, Jinan 250353, China

5. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

6. Shandong HiCon New Media Research Institute Co., Ltd., Jinan 250013, China

Abstract: Human activity sensing based on Wi-Fi channel state information (CSI) has an important application prospect in future intelligent interaction scenarios such as virtual reality, intelligent games, and the metaverse. Accurate sensing of complex and continuous human activities is an important challenge for Wi-Fi sensing. Convolutional neural network (CNN) has the ability of spatial feature extraction but is poor at modeling the temporal features of the data. While long short-term memory (LSTM) network or gated recurrent unit (GRU) network, which are suitable for modeling time-series data, neglect learning spatial features of data. In order to solve this problem, an improved CNN that integrates bidirectional gated recurrent unit (BiGRU) network was proposed. The bi-directional feature extraction ability of BiGRU was used to capture the correlation and dependence of the front and back information in the time series data. The extraction of the spatiotemporal features of the time series CSI data was realized, and then the mapping relationship between the action and the CSI data was present. Thus the recognition accuracy of the complex continuous action was improved. The proposed network structure was tested with basketball actions. The results show that the recognition accuracy of this method is above 95% under various conditions. Compared with the traditional multi-layer perceptron (MLP), CNN, LSTM, GRU, and attention based bidirectional long short-term memory (ABLSTM) baseline methods, the recognition accuracy has been improved by 1%~20%.

Key words: channel state information, human activity sensing, complex continuous action, convolutional neural network, bidirectional gated recurrent unit

0 引言

随着虚拟现实 (VR, virtual reality)^[1]、智能游戏^[2]、元宇宙^[3]等新一代信息技术应用场景的不断发展,智能人机交互需求不断增加。人体活动感知作为连接物理世界和信息世界的重要枢纽,是人机交互中最为重要的技术,在智能家居^[4]、辅助医疗^[5]、安防监控^[6]、情感计算^[7]等实际应用中的作用越来越重要。

人体活动感知技术主要分为基于可穿戴设备的感知技术、基于计算机视觉的感知技术和基于 Wi-Fi 信号的感知 (Wi-Fi 感知) 技术 3 种类型。其中,基于可穿戴设备的感知利用专用传感器感知用户活动,可以实现较高的识别精度^[8]。然而,这种方式需要用户佩戴一个甚至数个设备,在有些场合下应用不便且成本较高。基于计算机视觉的感知借助图像识别和视频分析技术,可以辨别人体细微的活动^[9]。但是,这种方式的性能受限于图像和视频的采集质量,对视线、光照和摄像机覆盖范围等应用条件的要求较高,且易泄露用户个人隐私,应用条件较为苛刻。基于 Wi-Fi 信号的感知利用泛在 Wi-Fi 基础设施获取信号数据,其核心原理是利用人体活动对 Wi-Fi 信号发射机与 Wi-Fi 信号接收机之间的无线信道传播特性所产生的相关变化实现

人体活动的感知与识别。Wi-Fi 感知无须用户佩戴额外设备,具有价格低廉、感知灵敏、隐私性高等优点,并且不受墙体、物体等环境限制或实际场景中光照条件的影响,因而在多种室内外人体活动感知场景中具有广泛的应用前景。

根据所利用的信号特征的不同,Wi-Fi 感知主要分为两类:基于接收信号强度 (RSSI, received signal strength indicator) 的感知和基于信道状态信息 (CSI, channel state information) 的感知。其中,基于 RSSI 的感知主要利用接收 Wi-Fi 信号的总体功率强度及其时间上的变化特征来对用户活动行为进行分析。由于功率信号强度模式较为单一,基于 RSSI 的感知一般仅适用于室内定位和简单人体活动识别。与基于 RSSI 的方法考虑功率和的方式不同,基于 CSI 的感知利用 Wi-Fi 正交频分复用 (OFDM, orthogonal frequency division multiplexing) 信道上的多个子载波的振幅和相位。与 RSSI 的单一的功率特征相比较,CSI 利用信号的时频域联合特征,能够表征更加丰富的信号特征,从而具有感知复杂人体活动的潜力^[10]。

目前,基于 Wi-Fi CSI 的人体活动感知主要采用机器学习或者深度学习算法对所采集的信号进行处理以识别人体活动特征^[11-13]。其中,机器学习算法一般由研究人员结合领域知识对信号的统计

特征进行分析与提取。例如,文献[2]提取了 CSI 数据的平均值、方差等统计特征实现活动的识别,文献[12]提取出手势动作的标准差、中位数绝对偏差、最大值、第 25 百分位数等特征作为最终的分类依据。但是,受限于提取效率和信号时变特性,人工提取特征只适用于特定的简单任务,普适性较差,难以泛化应用于复杂连续动作的识别。深度学习利用深度神经网络(DNN, deep neural network)、卷积神经网络(CNN, convolutional neural network)、递归神经网络(RNN, recurrent neural network)能够直接从输入的 CSI 中进行自动学习和自动特征提取,避免了传统机器学习算法特征定义和参数优化的复杂过程,近年来成为 Wi-Fi 感知技术发展的主流^[14]。文献[15]基于一维 CNN,通过对原始 CSI 时序数据的特征提取实现手部敲击键盘动作的识别。文献[16]将 CSI 数据转化为热图,通过二维 CNN 提取图像特征实现人员行为识别。文献[10,14]通过长短期记忆(LSTM, long short-term memory)和门控循环单元(GRU, gated recurrent unit)提取动作序列的时序特征,从而实现活动识别与人员身份识别。在上述基于深度学习的感知技术中,研究人员或利用 CNN 提取动作空间局部相关特征,或结合具有时序信息建模能力的 LSTM、GRU 等网络结构对信号中的时间相关特征进行提取,从而实现人体活动特征的感知与识别。然而,这些现有方案一般直接应用 CNN、LSTM、GRU 等模块刻画所感知动作在空间维度或时间维度上的特征信息。虽然经典的 LSTM 和 GRU 结构能够捕捉时间序列数据中的长期依赖关系,提取动作数据的时序特征,但是,它们仅考虑了与动作过去信息之间的联系,忽视了对动作未来信息的综合考量,无法全面提取动作数据的时序特征。不仅如此,上述 Wi-Fi 感知方法所识别的对象大多是跑、跳、走等比较简单的日常动作。能否利用 Wi-Fi 感知对振幅变化剧烈、活动频率较高的复杂连续活动(如体育活动、健身活动、极限运动等)进行高精度识别是值得进一步研究的问题。此外,上述方案大多仅采用 CSI 的振幅信息而忽视了 CSI 的相位信息,而相位信息对感知细粒度动作具有重要的价值^[17]。能否利用相位信息来提高复杂连续人体活动的识别精度值得进一步探索。

针对以上问题,本文提出一种新型的复杂连续人体活动 Wi-Fi 感知方法。所提方法的基本思想是

将 CNN 与 BiGRU 结合,构建起一种并行架构的 CNN-BiGRU。其中,CNN 用于提取动作的空间特征,BiGRU 用于学习动作的过去与未来的双向时序特征。相较于传统的 LSTM、GRU 等循环网络结构,BiGRU 利用双向(即过去信息与未来信息两个方向)特征提取结构能够处理时序信息和捕捉前后时序数据信息的关联性和依赖性。

所构建的 CNN-BiGRU 以预处理后的 CSI 振幅和相位为输入,通过并联方式的 CNN 和 BiGRU 自动学习和提取动作数据特征,并将融合后的时空特征作为动作的分类依据。以篮球动作为场景验证了 CNN-BiGRU 架构在复杂连续动作识别上的有效性。为验证不同环境对所提方法识别性能的影响,在不同场所的多种实验条件下进行实验,证明了该方法可以对复杂连续动作实现高精度识别。与文献[2,18]相比,本文利用深度学习的技术自动地提取动作的时空特征,实现了对复杂连续动作的有效识别,相较于经典的深度学习算法,识别准确率提升了 1%~20%。

本文的贡献可以概括如下。

1) 提出了一种基于 CNN-BiGRU 的架构,通过融合 BiGRU,实现了动作数据时空特征的提取,克服了单一结构对数据特征提取的限制,从而提高其分类准确率。

2) 区别于传统仅考虑振幅的设计,本文综合考虑了振幅和相位信息。实验表明,相位信息对复杂动作的识别性能具有提升作用。

3) 采集了 7 类篮球动作数据。以所采集的数据为基础,对所提网络结构进行了训练和验证,证明了 CNN-BiGRU 在复杂连续动作识别上的有效性。

1 相关工作

目前人体活动感知的相关研究工作主要分为基于可穿戴设备的感知、基于计算机视觉的感知以及基于 Wi-Fi 信号的感知。本节总结近年来的相关研究成果,并对 3 种方法加以比较。

1.1 基于可穿戴设备的感知

基于可穿戴设备的感知主要通过陀螺仪、加速度计、气压计等硬件设备获取活动数据^[19],对数据进行分析从而识别相应的活动。文献[20]开发了一种名为 BodyScope 的可穿戴声学传感器,BodyScope 可以记录用户喉咙区域产生的声音,并将其分类为相对应的用户活动,如吃饭、喝水、说

话、大笑和咳嗽。文献[21]使用嵌入多种传感器的智能手机,对走、跑和上下楼等基本动作进行分类。

1.2 基于计算机视觉的感知

与基于可穿戴设备的感知不同,基于计算机视觉的感知解决了用户需要佩戴传感器的问题,文献[22-24]的工作实现了对人体姿态和动作的自动检测和跟踪,它们利用摄像机监控人类活动,通过对图像或视频的捕获,实现对运动过程中各种参数的测量和分析,从而对活动进行感知。

1.3 基于 Wi-Fi 信号的感知

与上述两种技术相比,基于 Wi-Fi 信号的感知技术通过分析所接收到的信号特征实现环境感知,克服了佩戴传感器的不便,且不受物理空间的限制或视距、光照条件的影响。与此同时,基于 Wi-Fi 的系统成本低,易于部署,对活动及周围环境敏感,可以有效识别不同的人体活动。

早期的活动识别研究是利用 RSSI 实现的,文献[25]利用从智能手机获取的 RSSI,对 4 种手势的识别率达 72%,对 11 种手势的识别率达 51%。文献[26]提出了基于 Wi-Fi 信号的被动人体活动识别系统,通过 RSSI 建立指纹,采用基于分类树的融合算法,对部分活动的识别准确率可达 100%,平均准确率为 72.47%。虽然基于 RSSI 的方法实现简单,但是准确率较低,并不是一种有效的解决方案。

CSI 可以从一些商业设备中获得,如 Intel 5300 网卡^[27]、Atheros 系列网卡^[28]。文献[11]探索身体运动与 CSI 振幅信息之间的关联,利用 Wi-Fi 设备天线的相关性提取多普勒频移作为分类特征,使用极限学习机对 10 种动作进行分类,平均准确率为 94.2%。文献[29]设计了一种数据增强的基于 Wi-Fi 的人体活动识别系统。具体而言,该系统在每一个振幅子载波上使用短时傅里叶变换将 CSI 信号转换为时频图,并将其叠加起来,随后使用 8 种数据增强的方法扩充现有数据集,最后将扩充后的数据集送入所设计的密集 LSTM 网络模型进行分类训练。实验表明合成数据的性能提高了 34.6%,准确率达到 90%。文献[30]提出了一种基于 BiLSTM 网络的手势识别算法,该算法提取原始数据中的振幅信息,通过主成分分析(PCA, principal component analysis)与低通滤波去除振幅噪声,通过所设计的 BiLSTM 架构实现 6 种手势的识别,平均识别准确率达 91.49%。

现有的研究工作在处理原始 CSI 数据时,往

往只提取振幅信息,而忽略相位的使用,导致一些特征信息缺失;所识别的动作绝大部分为日常动作,动作相对简单。本文将振幅与相位相结合,作为神经网络的输入,确保 CSI 的数据完整性。并且,所针对的篮球动作是全身性的复杂连续动作,识别难度更大。篮球动作的识别工作^[31-32]已通过计算机视觉技术或者利用传感器实现,它们为人机交互或 VR 技术的发展提供了一定的帮助。本文利用 Wi-Fi 感知技术对篮球动作进行识别,成本低廉、鲁棒性高。

2 Wi-Fi 感知系统模型

Wi-Fi 感知系统总体架构如图 1 所示,由数据采集模块、数据预处理模块以及活动分类模块 3 个模块组成。其中,数据采集模块通过网卡获取原始 CSI 数据;数据预处理模块对所采集的原始 CSI 数据中的振幅衰减和相位偏移进行降噪和校准,并基于振幅方差将 CSI 数据划分为活动部分与非活动部分;活动分类模块是由神经网络组成的分类器,它以降噪和校准后的活动部分 CSI 数据为输入,对 CSI 时空信息进行自动的特征提取及分类。

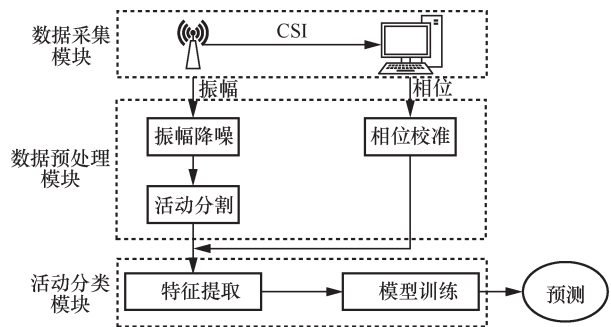


图 1 Wi-Fi 感知系统总体架构

2.1 基于 CSI 的 Wi-Fi 感知原理

现有商业化 Wi-Fi 技术大多采用 IEEE 802.11 a/g/n 无线通信协议,其核心是利用 OFDM 将信号调制到多个子载波上并行传输^[33]。OFDM 的本质是将宽带信道转换为多个并行窄带信道,每个子载波上的信道可以看作水平衰落信道,从而大大降低了接收机均衡器的复杂度。OFDM 下变频后的基带信号可以表示为

$$Y = HX + N \quad (1)$$

其中, $H \in \mathbf{C}^{K \times T}$ 为 CSI 矩阵, $X \in \mathbf{C}^{K \times T}$ 和 $Y \in \mathbf{C}^{K \times T}$ 分别为发送和接收的基带信号矩阵, N 为传输过程中的噪声矩阵。 K 和 T 分别表示子载波和时隙数目。

OFDM 的 CSI 矩阵 \mathbf{H} 取决于无线信号传播的环境。无线信号在传播过程中所经历的路径损耗、反射、散射和折射等环境因素均会影响 OFDM 时频子载波信道参数, 即 CSI。也就是说, 物理空间中运动的物体会对 CSI 的时频特征产生时变影响。本文实际采集的原始 CSI 振幅三维图如图 2 所示。其中包索引对应时间域、子载波索引对应频率域。所谓 Wi-Fi 感知就是通过对 Wi-Fi 接收设备所获取的 OFDM CSI 的时频特征进行分析, 从而识别人体活动的类型。

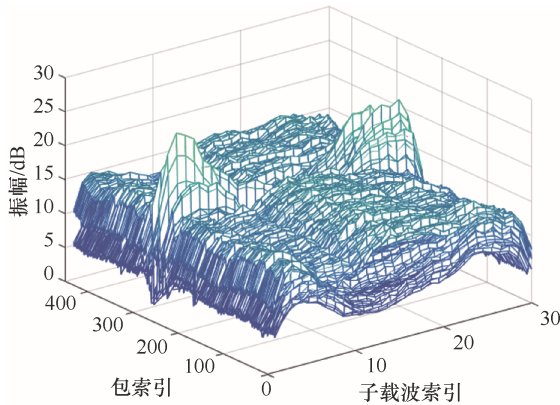


图2 原始 CSI 振幅三维图

2.2 CSI 数据的获取

Wi-Fi CSI 可以通过对接收端的 Wi-Fi 网卡数据进行解析得到。常见的能够方便获取 CSI 的网卡或设备包括 USRP、Intel 5300 网卡和 Atheros 系列网卡等。本文采用 1 台装有 3 根天线的 Intel 5300 网卡的 PC 接收由 TP-Link WR742N Wi-Fi 路由器发送的 Wi-Fi 信号, 并利用 CSI Tool^[27] 解析获得 3 路 CSI。每根天线的 CSI 都具有 $K=30$ 个 OFDM 子载波。对于第 m 根天线, 在 T 个连续时隙上可以获得一个 CSI 矩阵, 表示为 $\mathbf{H}_m \in \mathbf{C}^{K \times T}$, $m=1,2,3$ 。

2.3 数据预处理模块

由于周围环境、信号干扰以及硬件设备的问题, 采集的 CSI 数据除了包含活动信号, 还包含噪声及异常值。为了提高信号质量, 需要对异常值和噪声进行预处理。

2.3.1 异常值去除

环境干扰、设备异常等原因造成的异常值, 其振幅与相邻数据存在显著差别。为了去除这些异常值的影响, 在振幅数据的每一个子载波上应用带有滑动窗口的 Hampel 滤波器。Hampel 滤波器将在 $[\mu - \gamma\sigma, \mu + \gamma\sigma]$ 外的点视为异常值, 而后再用 μ 替换异常值。其中 μ 和 σ 分别为滑动窗口内数据序列

的中位数和中位数绝对偏差, γ 为常数因子, 通常设为 3^[34]。经 Hampel 滤波后, 可以有效地去除异常值。

2.3.2 噪声去除

去除异常值后, 进一步采用低通滤波器对信号进行滤波以降低 CSI 序列中的噪声。本文选择 Savitzky-Golay 滤波器^[35]对 CSI 时间序列进行平滑。具体而言, 令 $x_t, t \in N$ 表示一个时间序列, 对以 x_t 为中心左右各 M 个样本点的数据进行窗口操作, 即取出以 x_t 为中心的 $(2M+1)$ 个采样点。利用 p 阶多项式对这些采样点进行拟合, 即

$$y_t = \sum_{k=0}^p a_k t^k, -M \leq t \leq M, p \leq 2M+1 \quad (2)$$

定义拟合的均方差函数为

$$Q = \sum_{t=-M}^M |y_t - x_t|^2 \quad (3)$$

其中, 参数 a_k 取使得式(3)最小的值。

在实际应用时, 在每一个 CSI 的子载波序列上应用 Savitzky-Golay 滤波器, 从而实现 CSI 数据的降噪。

2.3.3 相位校准

原始 CSI 数据中提取出的相位信息存在载波频率偏移 (CFO, carrier frequency offset) 和采样频率偏移 (SFO, sample frequency offset), 无法直接使用。为此, 利用线性变换^[36]的方法对相位信息进行校准。

所获得的第 i 个子载波上的原始相位表示为

$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N} \Delta t + \beta + z \quad (4)$$

其中, $\hat{\phi}_i$ 为原始相位, ϕ_i 为真实相位, Δt 是 SFO 引起的时间偏差, β 是 CFO 引起的未知相位偏差, z 为测量噪声, k_i 表示第 i 个子载波的索引, N 是快速傅里叶变换的长度 (IEEE 802.11n 中 $N=64$)。接下来, 通过从原始相位中减去线性项 $ak_i + b$, 即可消除 Δt 和 β , 得到校准后的相位。其中, 线性项中 $a = \frac{\hat{\phi}_n - \hat{\phi}_1}{k_n - k_1}$, $b = \frac{1}{n} \sum_{j=1}^n \hat{\phi}_j$, 经线性变换校准后的相位可表示为

$$\tilde{\phi}_i = \hat{\phi}_i - ak_i - b = \phi_i - \frac{\phi_n - \phi_1}{k_n - k_1} k_i - \frac{1}{n} \sum_{j=1}^n \phi_j \quad (5)$$

数据预处理前后的振幅和相位对比如图 3 所示 (以投篮动作为例), 所采用的数据预处理方法

实现了对振幅的降噪,并使相位成为可用于检测的信号。

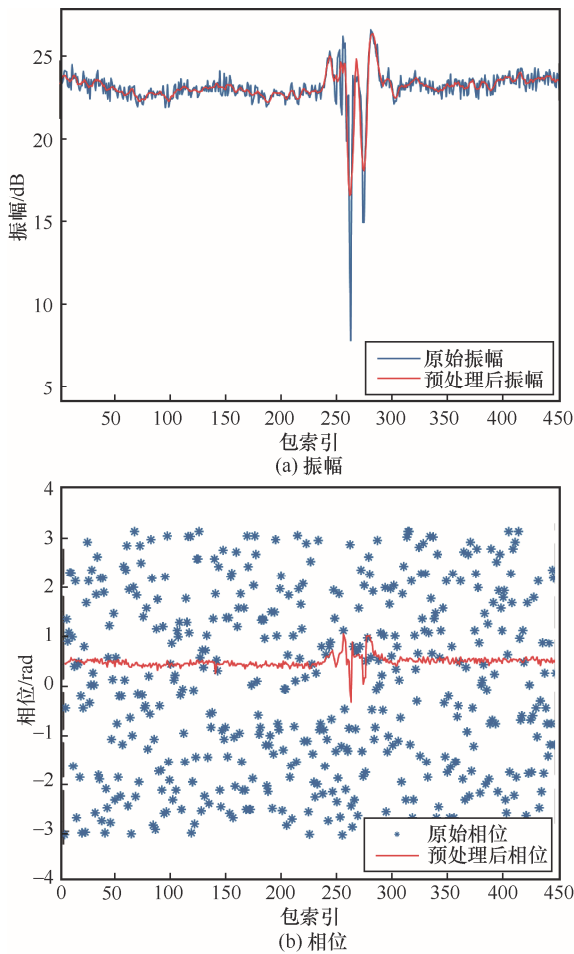
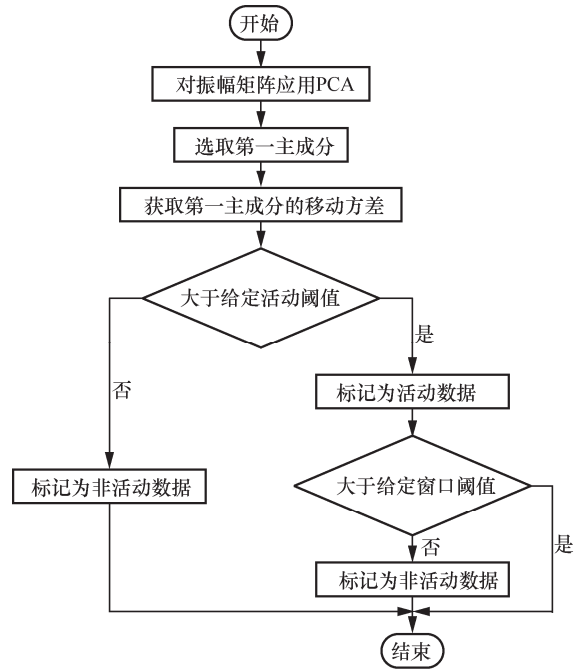


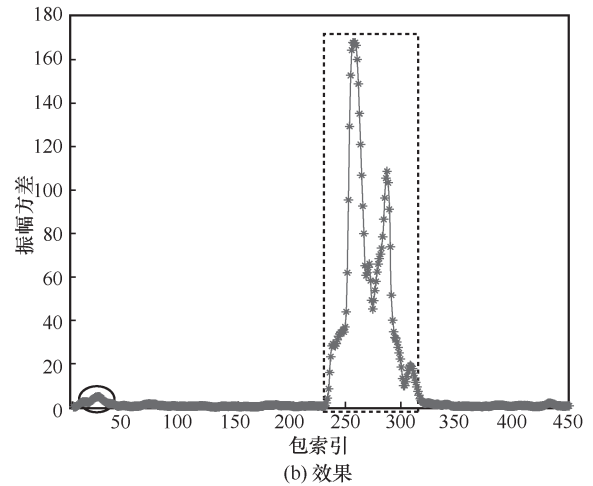
图3 数据预处理前后的振幅和相对比

2.3.4 活动分割

由于 CSI 数据中包含静态信息,即非活动的部分,如果将此部分 CSI 数据输入神经网络,会增加算法的复杂度。因此需要有效区分活动部分与非活动部分的数据,丢弃非活动部分的数据,将活动部分的振幅和相位信息作为神经网络的输入,完整的活动数据也是提高神经网络分类准确率的关键。对于复杂剧烈的连续活动,活动部分数据的方差远大于非活动部分数据的方差,因此,基于此现象预先设置活动阈值 ζ 。此外,由于 Wi-Fi 信号的敏感性,在非活动部分可能会存在短暂的波动,这些波动可能会被错误地归类为活动部分,为了获得更加完整且准确的活动数据,引入窗口阈值 η ,目的是去除该波动带来的错误划分。活动分割如图 4 所示,流程如图 4(a)所示,给出本文所提基于双阈值的活动分割方法的具体步骤。



(a) 流程



(b) 效果

图4 活动分割

步骤 1 对振幅所构成的矩阵应用 PCA, 自动选择代表 CSI 时间序列中最常见变化的主成分, 得到其主成分矩阵, 主成分矩阵反映了子载波振幅的变化。

步骤 2 选择第一主成分进行活动分割, 通过应用滑动窗口的方式, 计算窗口内数据点的方差, 返回方差所构成的数据序列, 以此获得第一主成分的移动方差, 作为活动分割的指标。

步骤 3 给定活动阈值 ζ , 当第一主成分的方差大于活动阈值 ζ 时, 活动开始; 当第一主成分的方差小于活动阈值 ζ 时, 活动结束。标记方差大于 ζ 的样本点, 作为 CSI 数据的活动部分。

步骤 4 引入窗口阈值 η , 再次将窗口大小

(即包索引) 小于 η 的样本点标记为非活动数据, 即可得到最终的标记数据。效果如图 4(b)所示, 虚线框中大致描述了活动部分的索引点标记, 实线圈内则是短暂的波动。

2.4 并行 CNN-BiGRU

近年来, LSTM^[10]、GRU^[14]、CNN^[37]等深度学习模型已被成功应用至基于 CSI 的活动识别。CSI 中不仅包含动作的空间特征, 还包含动作的时序特征。LSTM、GRU 能够学习长序列信息之间的依赖和关联关系, 记录间隔或时延较大的历史信息 and 重要事件^[30]。然而, 具有时序建模能力的 LSTM、GRU 网络并未考虑动作空间特征的提取。CNN 具有局部链接、权重共享的特点, 但具备强大的特征提取能力的 CNN 却忽视了时序信息之间的相关性。此外, LSTM 和 GRU 因时序信息只能单个

方向传递的结构特点, 仅能考虑动作过去的时序信息, 忽略了对未来信息的规律学习。因此, 现有工作中活动识别系统并未充分挖掘动作的时空特征, 从而不能达到较高的识别精度。为此, 本文提出一种融合 CNN 与 BiGRU 的新型 Wi-Fi 人体活动感知架构。

CNN-BiGRU 架构如图 5 所示, 输入信号是由 Wi-Fi 网卡的 3 根天线对应的 CSI 矩阵 $H_m (m=1,2,3)$ 的振幅部分和相位部分堆叠扩展后的二维矩阵

$$\bar{H} = [|H_1|, |H_2|, |H_3|, \angle H_1, \angle H_2, \angle H_3]_{2MK \times T} \quad (6)$$

其中, $|H_m|$ 和 $\angle H_m$ 分别表示对矩阵 H_m 中每个元素提取振幅和相位后构成的振幅矩阵和相位矩阵, $M=3$ 表示共有 3 根天线的数据。

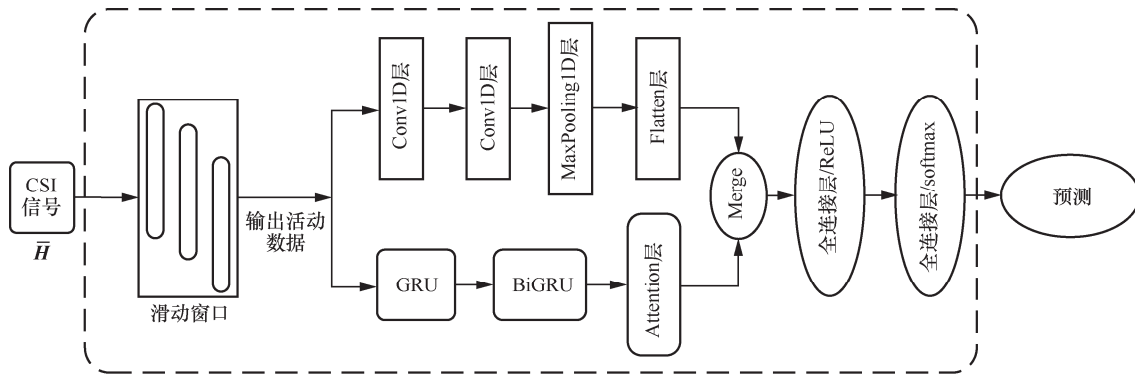


图 5 CNN-BiGRU 架构

神经网络要求输入的数据维度一致。但是, 由于每个动作的持续时间不尽相同, 其数据包的长短是不同的。另外, 输入的数据维度增大, 算法的时间复杂度也会增加。因此, 为了使输入数据维度一致, 并且降低算法的复杂度, 所设计的网络在输入层应用滑动窗口沿时间序列方向对二维矩阵进行分割, 并将标记活动部分样本点少于 60% 的分割片段丢弃以去除 CSI 中的非活动数据, 获得相同维度的数据片段, 保留的分割后的数据片段作为网络的最终输入。输入的数据经两个分支进行特征的提取, 特征融合, 作为最后分类的依据。分支一基于一维 CNN 构建, 用于提取篮球动作空间维度上的特征。分支二基于 GRU 和 BiGRU 构建, 用于提取时间维度上的特征。综合所提取空间维度和时间维度的特征, 作为最后的分类依据, 通过 softmax 函数对动作进行分类。

2.4.1 CNN

本文所采用的 CNN 结构如图 6 所示。本文利用 CNN 的局部连接、权重共享等特性, 提取出预处理后 CSI 数据的高层特征。

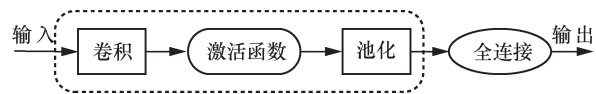


图 6 CNN 结构

2.4.2 GRU 与 BiGRU

RNN 具有短期记忆的能力, 在处理短时时序序列的问题上有显著优势。但是, 若所处理的时序序列维数过高, 会存在梯度消失的问题。LSTM 和 GRU 的相继提出改善了此问题。

GRU 是 LSTM 的高级变体, 相较于 LSTM 简化了门控机制, 并且不引入额外的记忆单元, 仅通过更新门和重置门控制信息的更迭。GRU

结构如图 7 所示，包含 3 个参数：更新门 z_t 、重置门 r_t 以及隐状态 h_t ，并通过式(7)~式(10)进行更新。

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \quad (7)$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \quad (8)$$

$$\tilde{h}_t = \tanh(W[h_r h_{t-1}, x_t]) \quad (9)$$

$$h_t = (1 - z_t)h_{t-1} + z_t\tilde{h}_t \quad (10)$$

其中， W_z 、 W_r 、 W 为权重矩阵， x_t 为 t 时刻的时序信息， h_{t-1} 表示 $(t-1)$ 时刻的隐藏状态， σ 为 sigmoid 激活函数。

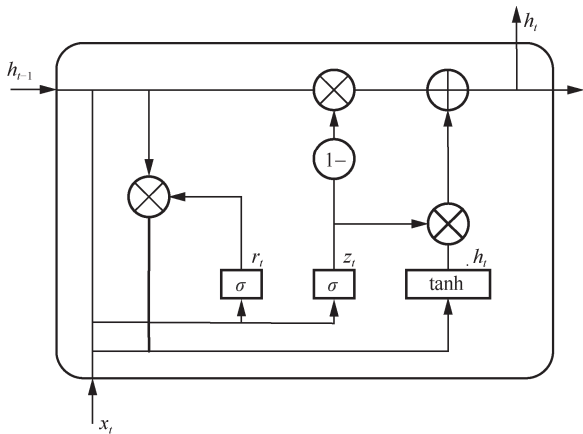


图 7 GRU 结构

由于所收集的篮球动作是连续性动作，过去的信息以及未来的信息对动作识别的影响同等重要。BiGRU 可以从过去以及未来两个方向进行时序特征的提取，因此选用 BiGRU 学习动作特征的双向规律，以提取更为全面的特征。双向 GRU 结构如图 8 所示，BiGRU 包含正向和反向 GRU，最终的状态 h_t 由正向和反向 GRU 的隐状态共同决定，并将此状态作为 BiGRU 的输出。

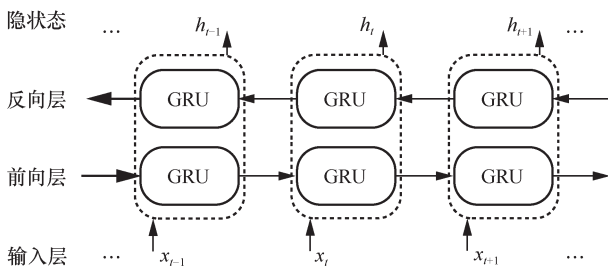


图 8 双向 GRU 结构

2.4.3 注意力机制

注意力机制最初是为机器翻译设计的，并在图像处理领域、自然语言处理领域广泛应用。可以用人类视觉机制进行直观解释，当某个人通过视觉感知某些事物时，通常情况下是根据需要关注某个感兴趣的区域，这样在将来出现类似的场景时，他将学会把注意力放在该区域^[38]。BiGRU 为所有 CSI 信号的特征分配相同的权重，而不同的特征对于篮球动作的识别可能有不同的贡献。因此，采用注意力机制为更重要的特征分配更高的权重，加强关键信息的影响，从而提高网络的识别性能。

注意力机制如图 9 所示，注意模型的输入为从 BiGRU 网络学习到的序列特征，记为 h_t ，其中， $1 \leq t \leq n$ ，通过 tanh 函数计算每个特征向量的重要性得到其得分 g_t ，表示为

$$g_t = \tanh(W^T h_t + b) \quad (11)$$

其中， W^T 为权重向量， b 为偏置。随后使用 $z_t = \text{softmax}(g_t)$ 对其得分进行归一化，最后将特征向量与归一化后的得分的乘积作为注意机制的最终输出，表示为

$$O = \sum_{t=1}^n z_t h_t \quad (12)$$

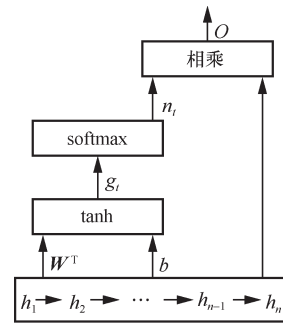


图 9 注意力机制

3 实验及结果分析

为了验证所提方法的有效性，本文进行了多组对比实验，实验场景及实验内容如下。

3.1 数据采集

发射端为单天线 TP-Link WR742N 路由器，工作在 IEEE 802.11n 协议的 2.4 GHz 频段，每秒传输 50 个数据包（即采样率为 50 Hz）。接收端为配备 8 dBi 增益全向天线的三天线 Intel 5300 网卡的台式

计算机，操作系统为 Ubuntu 14.04，通过 CSI Tool 采集 3 路 CSI 数据。每路数据对应 1 根天线，具有 30 个子载波。

在办公室和实验室两个环境中进行数据的采集工作，以对所提方法进行评估，实验场景如图 10 所示，其中 TX 为发射端，RX 为接收端，测试人员在 TX 与 RX 间完成相应的动作。招募了 5 名篮球爱好者（4 男 1 女），每名测试者在规定时间内完成篮球动作的演示，活动的开始与结束时刻测试者保持静止，以减少非篮球动作的干扰，确保数据的准确性。

实验设计了 7 种常见篮球动作，每种动作的具体描述如下。

(1) 立定投篮：测试者双脚分开与肩同宽，屈膝后向上模拟投出动作，当手臂垂直后手腕模拟拨腕。

(2) 跳投：测试者双脚分开与肩同宽，屈膝后向上跳起，并做出投篮的动作，在跳起最高点模拟拨腕。

(3) 运球：测试者做出半蹲动作，非运球手做出护球动作，运球手模拟两次拍球。

(4) 上篮：测试者模拟三步上篮动作，前两步为向前跨步，在第三步起跳并模拟拨腕。

(5) 盖帽：测试人员屈膝后向上跳起，在最高处模拟盖帽动作。

(6) 抢断：测试者做出半蹲动作，朝抢断手方向俯冲，模拟出断球动作。

(7) 转身：测试者屈膝后，向右做出转身动作。

每种实验场景下的每种篮球动作共做 50 次，共收集 2 800 个动作数据。

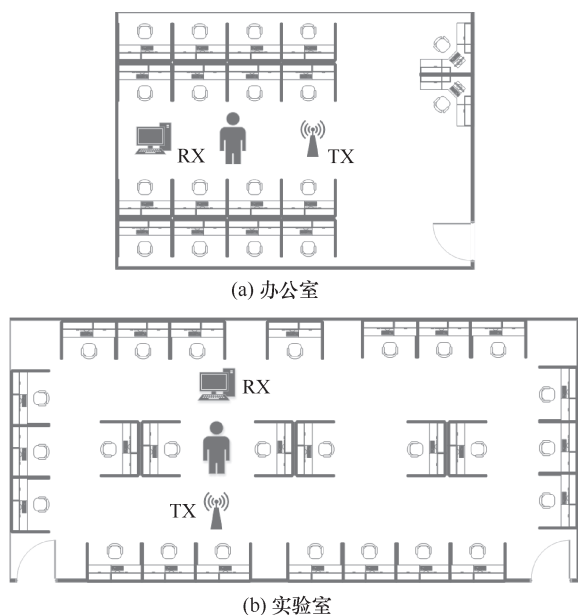


图 10 实验场景

3.2 模型及训练参数设置

本文以 7:2:1 的比例将数据集划分为训练集、验证集和测试集，模型详细参数设置见表 1，实验环境配置数据见表 2。训练集以及验证集识别准确率变化曲线如图 11 所示。

表 1 模型详细参数设置

| 参数 | 设置 |
|------------|---------|
| 滑动窗口大小 | 0.2 s |
| 卷积核大小 | 3 |
| 卷积核数量 | 128、256 |
| GRU 隐藏单元 | 200 |
| BiGRU 隐藏单元 | 400 |
| 注意力机制隐藏单元 | 400 |
| 全连接层参数 | 128、7 |

表 2 实验环境配置数据

| 参数 | 设置 |
|-----------|-------------------|
| CPU | AMD Ryzen 7 4800H |
| 显卡 | GeForce GTX 1650 |
| 深度学习框架 | TensorFlow |
| 优化器 | Adam |
| 损失函数 | 交叉熵 |
| 学习率 | 0.000 03 |
| Epochs | 300 |
| BatchSize | 64 |

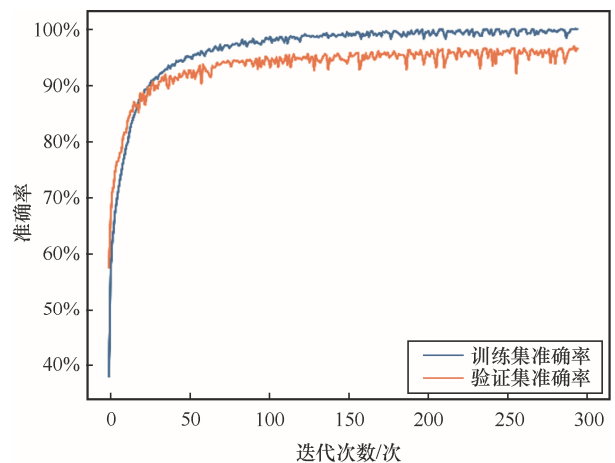


图 11 训练集以及验证集识别准确率变化曲线

3.3 实验评估

本文采取不同的评估指标验证所提方法的可靠性及有效性，包括准确率、精确率、召回率和 F1 值^[12]。

3.3.1 WiAR 数据集实验

本文首先在公开数据集 WiAR^[39]进行了实验，以证明所提方法的有效性。实验提取了其常见的 7 种日常动作进行了验证，分别为水平挥臂、垂直挥臂、鞠躬、前踢、走、坐下和蹲起动作。

实验选取了 Wi-Fi 感知领域中经典的深度学习算法进行比较。文献[40]通过构建 MLP 实现对人员的计数。文献[15]基于 CNN 构建了分类模型，实现了敲击键盘动作的分类。文献[10]、文献[14]和文献[41]分别基于 LSTM、GRU 和 BiLSTM 构建神经网络，实现了活动识别与身份识别。实验结果见表 3，本文所提基于 CNN-BiGRU 的方法综合提取了动作的时空特征，在 WiAR 数据集上取得了最高的平均识别准确率。

3.3.2 环境实验

本文在两个实验场景中进行了实验，办公室场景较为空旷，实验室场景更为复杂，以此验证所提方法的鲁棒性。

混淆矩阵如图 12 所示，图 12(a)、图 12(b)分别展示了两个实验场景的混淆矩阵，其中每一列代表篮球动作的预测类别，每一行代表篮球动作的真实类别

类别。由混淆矩阵可知所提方法可以非常准确地对 7 种篮球动作进行有效分类，在办公室与实验室场景下的平均识别准确率分别为 97.4%、97.1%。同时还制作了包含每个动作的精确率、召回率以及 F1 值相关柱状图，以进行综合评价。办公室和实验室场景下每个活动的精确率、召回率和 F1 值如图 13 所示，所有活动均在合理范围内分类。

3.3.3 距离实验

为验证办公室环境中收发机距离对所提方法识别性能的影响，分别设置收发机距离为 2 m、3 m、4 m 和 5 m，不同距离实验结果如图 14 所示。可以看出所提方法在不同的收发机距离条件下，整体性能保持稳定。收发机距离为 2 m 时获得了最高的识别准确率；在 5 m 时识别准确率最低，但也达到了高于 95% 的识别准确率。一般情况下，随着设备距离的增加，识别准确率会有所降低，而在 4 m 时出现了略高于 3 m 的情况，猜测是数据采集过程中篮球动作的剧烈性导致某些数据不稳定。综合而论，在室内环境下，所提方法在不同距离条件下能保持较高的准确率。

表 3 实验结果

| 方法 | 平均识别准确率 | 水平挥臂 | 垂直挥臂 | 鞠躬 | 前踢 | 走 | 坐下 | 蹲起 |
|------------------------|---------|-------|-------|-------|-------|-------|-------|-------|
| MLP ^[40] | 76.1% | 86.0% | 82.0% | 65.0% | 53.0% | 93.0% | 87.0% | 67.0% |
| CNN ^[15] | 88.7% | 96.0% | 94.0% | 84.0% | 71.0% | 97.0% | 90.0% | 89.0% |
| LSTM ^[10] | 91.1% | 96.0% | 95.0% | 86.0% | 78.0% | 97.0% | 92.0% | 94.0% |
| GRU ^[14] | 92.6% | 98.0% | 95.0% | 85.0% | 83.0% | 98.0% | 94.0% | 95.0% |
| ABLSTM ^[41] | 94.0% | 98.0% | 95.0% | 89.0% | 90.0% | 98.0% | 95.0% | 93.0% |
| CNN-BiGRU | 96.0% | 98.0% | 97.0% | 94.0% | 92.0% | 99.0% | 97.0% | 95.0% |

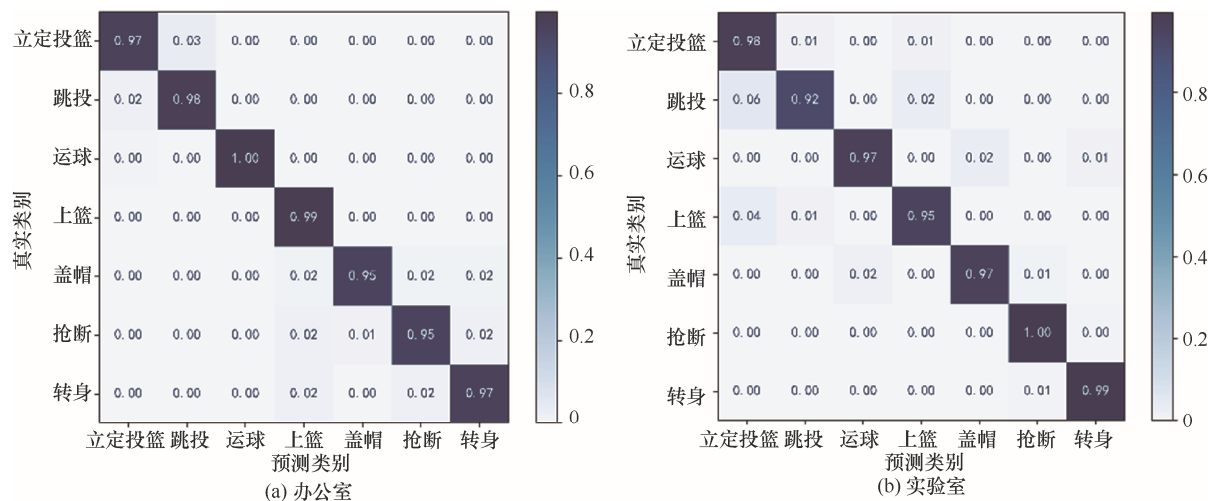
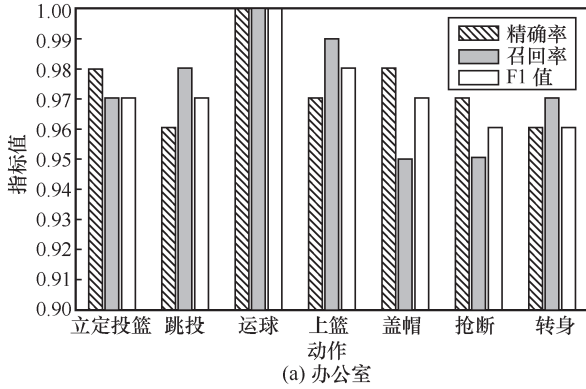
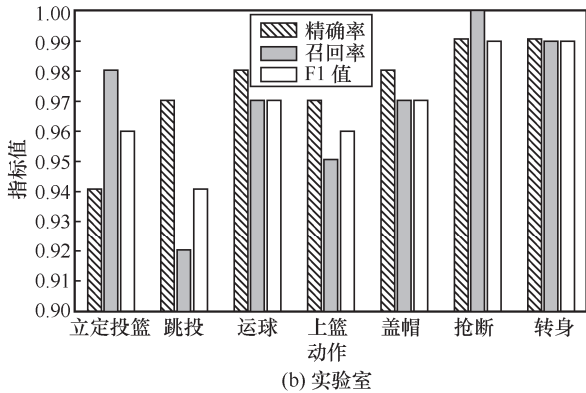


图 12 混淆矩阵



(a) 办公室



(b) 实验室

图 13 精确率、召回率和 F1 值

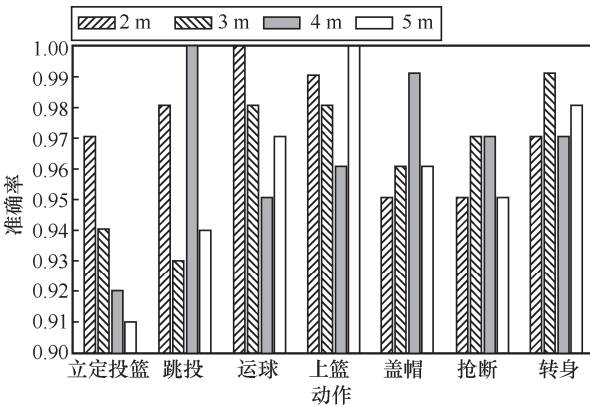


图 14 不同距离实验结果

3.3.4 测试者位置实验

为了验证办公室环境中测试者所处不同位置对识别性能的影响，在保持收发机距离为 2 m 的前提下，在 3 个不同的位置进行实验，不同测试者位置如图 15 所示，位置 1 在收发机连线的中点位置，位置 2、位置 3 分别为从位置 1 出发，垂直于收发机连线向右、向左前进 1 m 的位置。不同测试者位置实验结果如图 16 所示，展示了 3 个位置的平均识别准确率，3 个位置的准确率相差 1% 左右，验证了所提方法可适应不同位置所做的动作，具有较高的准确率。

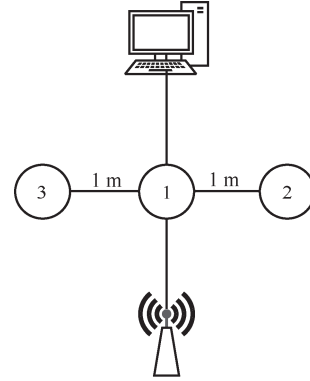


图 15 不同测试者位置

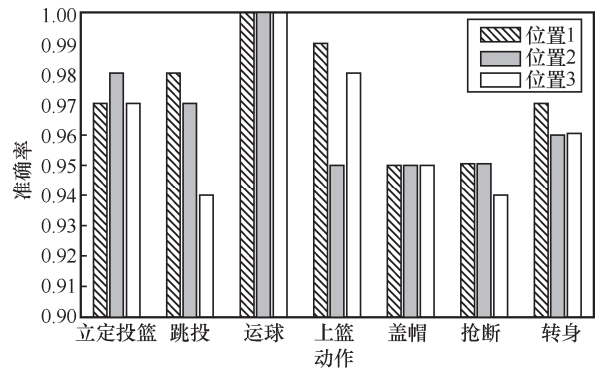


图 16 不同测试者位置实验结果

3.3.5 深度学习算法实验

本文采用第 3.3.1 节中提到的 5 种基于深度学习的感知算法进行了对比实验，不同深度学习算法实验结果如图 17 所示。所设计的 CNN-BiGRU 综合考虑篮球动作的时空特征，建立了 CSI 数据与篮球动作之间的映射关系，在与 5 种经典深度学习算法对比中取得了最高的识别准确率。

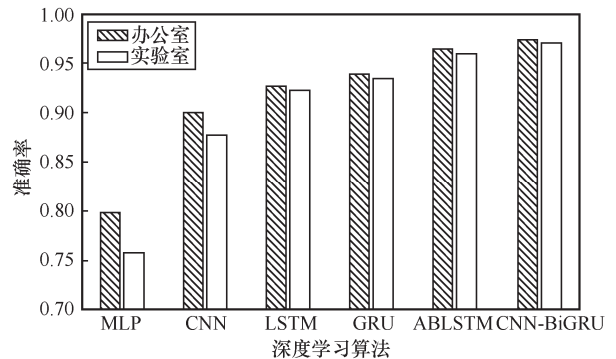


图 17 不同深度学习算法实验结果

3.3.6 时间开销分析

本文记录了每种深度学习算法的时间成本，即训练和测试的时间成本之和，不同深度学习算法训练和测试成本见表 4。虽然所提方法训练和测试时间

成本最大，但是获得了高于经典算法 1%~20%的识别准确率。此外，经实验测试，所提 CNN-BiGRU 结构的测试时间小于 5 s，因此，所提方法在具有较高识别准确率的同时也具备一定的实时性。

表 4 不同深度学习算法训练和测试成本

| 模型 | 时间开销/s |
|-----------|--------|
| MLP | 138 |
| CNN | 180 |
| LSTM | 228 |
| GRU | 210 |
| ABLSTM | 384 |
| CNN-BiGRU | 540 |

3.3.7 消融实验

本文评估了方法框架中每一部分对识别性能

的影响，即基信号、数据预处理以及网络架构，办公室、实验室环境中不同模块的平均识别准确率分别见表 5、表 6，表中“√”表示实验包含了该部分。

首先，对 3 种基信号的选择进行了实验，在没有经过预处理的情况下，仅使用振幅信息获得的准确率要高于振幅与相位的组合。这是由于未经相位净化的相位信息存在严重的随机相位偏移，使其不适用于活动的判别，从而影响总体识别准确率。

其次，选择振幅与相位的组合作为基信号，并分别对其进行降噪和校准处理时，在两个环境中达到了最高的识别准确率。

最后，本文测试了 CNN-BiGRU 中模型分支一和模型分支二对识别性能的影响，分别使用分支一、分支二进行模型训练时，仍可以达到较高的识别准确率，这也证明了本文网络结构选择的有效性。将

表 5 办公室环境中不同模块的平均识别准确率

| 实验序号 | 是否含有此模块 | | | | | | | 平均识别准确率 | |
|------|---------|----|-------|-------|------|------|-----|---------|-----------|
| | 基信号 | | | 数据预处理 | | 网络架构 | 分支二 | | |
| | 振幅 | 相位 | 振幅+相位 | 振幅降噪 | 相位校准 | | | | 分支一（卷积模块） |
| 实验 1 | √ | | | | | √ | √ | √ | 96.1% |
| 实验 2 | | √ | | | | √ | √ | √ | 62.5% |
| 实验 3 | | | √ | | | √ | √ | √ | 95.4% |
| 实验 4 | √ | | | √ | | √ | √ | √ | 96.8% |
| 实验 5 | | √ | | | | √ | √ | √ | 90.1% |
| 实验 6 | | | √ | √ | √ | √ | √ | √ | 97.4% |
| 实验 7 | | | √ | √ | √ | | √ | √ | 96.6% |
| 实验 8 | | | √ | √ | √ | √ | | | 95.6% |
| 实验 9 | | | √ | √ | √ | √ | √ | | 96.8% |

表 6 实验室环境中不同模块的平均识别准确率

| 实验序号 | 是否含有此模块 | | | | | | | 平均识别准确率 | |
|------|---------|----|-------|-------|------|------|-----|---------|-----------|
| | 基信号 | | | 数据预处理 | | 网络架构 | 分支二 | | |
| | 振幅 | 相位 | 振幅+相位 | 振幅降噪 | 相位校准 | | | | 分支一（卷积模块） |
| 实验 1 | √ | | | | | √ | √ | √ | 95.9% |
| 实验 2 | | √ | | | | √ | √ | √ | 56.8% |
| 实验 3 | | | √ | | | √ | √ | √ | 95.5% |
| 实验 4 | √ | | | √ | | √ | √ | √ | 96.6% |
| 实验 5 | | √ | | | | √ | √ | √ | 88.4% |
| 实验 6 | | | √ | √ | √ | √ | √ | √ | 97.1% |
| 实验 7 | | | √ | √ | √ | | √ | √ | 96.2% |
| 实验 8 | | | √ | √ | √ | √ | | | 95.0% |
| 实验 9 | | | √ | √ | √ | √ | √ | | 96.7% |

两个分支并行提取特征，训练性能有所提升（去除了注意力机制）。在此基础上加入注意力机制后，性能得到了进一步的提升，达到了最佳的识别准确率。

3.3.8 真实场景实验

本文在真实场景中验证了所提方法的有效性。实验场地为一个篮球场馆，真实实验场景及实验结果如图 18 所示，实验场景如图 18(a)所示，测试人员在收发机之间面向篮筐模拟不同的篮球动作。实验结果如图 18(b)所示，展示了所提方法在真实场景中的混淆矩阵，每种动作的识别准确率均在 85%之上，7 种动作的平均识别准确率达 92.1%。与办公室和实验室场景相比，识别准确率有所下降，这主要是因为所在环境较为嘈杂（临近公路且有人的存在），信号受到的干扰更多，对识别准确率产生了一定的影响。综合而论，所提方法在真实场景中的识别准确率达 92.1%，验证了其在真实场景中的有效性。

4 结束语

本文提出了一种基于 CNN-BiGRU 的复杂连续人体活动 Wi-Fi 感知方法。它可以提取动作丰富的时空特征，实现对复杂且连续的人类动作的高精度识别。本文以篮球动作为例，采用深度学习技术证明了所提方法以及并行 CNN-BiGRU 对复杂连续活动识别的有效性。实验结果表明，所提方法在含有相似运动成分的动作分类中（如立定投篮和跳投，均含有屈膝、拨腕的动作成分；上篮、盖帽也均有类似的手部动作成分），取得了可观的分类精度，平均识别准确率高于 95%。当然，主要错分也多存在于这些包含相似运动成分的动作中。

未来的工作需要更深层次地挖掘动作样本中的信息，通过增加天线数目或者增加收发机的数量继续提高包含相似运动成分的动作的分类精度。此外，在未来的工作中，会进一步拓展已有研究，如双人交互动作的识别，以篮球动作为背景（如一人投篮另一人防守），探索 Wi-Fi CSI 与其之间的联系，为活动识别研究提供进一步的参考。

参考文献：

- [1] 张宇翔, 任爽. 定位技术在虚拟现实中的应用综述[J]. 计算机科学, 2021, 48(1): 308-318.
ZHANG Y X, REN S. Overview of application of positioning technology in virtual reality[J]. Computer Science, 2021, 48(1): 308-318.
- [2] CHEN C, SHU Y, SHU K I, et al. WiTT: modeling and the evaluation of table tennis actions based on Wi-Fi signals[C]//Proceedings of 2018 24th International Conference on Pattern Recognition (ICPR). Piscataway: IEEE Press, 2018: 3100-3107.
- [3] 王文喜, 周芳, 万月亮, 等. 元宇宙技术综述[J]. 工程科学学报, 2022, 44(4): 744-756.
WANG W X, ZHOU F, WAN Y L, et al. A survey of metaverse technology[J]. Chinese Journal of Engineering, 2022, 44(4): 744-756.
- [4] WANG Y, LIU J, CHEN Y Y, et al. E-eyes: device-free location-oriented activity identification using fine-grained Wi-Fi signatures[C]//Proceedings of the 20th annual international conference on Mobile computing and networking. New York: ACM Press, 2014: 617-628.
- [5] LI F, VALERO M, SHAHRIAR H, et al. Wi-COVID: a COVID-19 symptom detection and patient monitoring framework using Wi-Fi[J]. Smart Health, 2021(19): 100147.
- [6] CHEN M D, MA J, ZENG X M, et al. MD-alarm: a novel manpower detection method for ship bridge watchkeeping using Wi-Fi signals[J]. IEEE Transactions on Instrumentation and Measurement, 2022(71): 1-13.



(a) 实验场景



(b) 实验结果

图 18 真实实验场景及实验结果

- [7] GU Y, WANG Y T, LIU T, et al. EmoSense: computational intelligence driven emotion sensing via wireless channel data[J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2020, 4(3): 216-226.
- [8] LARA O D, LABRADOR M A. A survey on human activity recognition using wearable sensors[J]. *IEEE Communications Surveys & Tutorials*, 2013, 15(3): 1192-1209.
- [9] JALAL A, KIM Y H, KIM Y J. Robust human activity recognition from depth video using spatiotemporal multi-fused features[J]. *Pattern Recognition*, 2017(61): 295-308.
- [10] YOUSEFI S, NARUI H, DAYAL S, et al. A survey on behavior recognition using Wi-Fi channel state information[J]. *IEEE Communications Magazine*, 2017, 55(10): 98-104.
- [11] YAN H, ZHANG Y, WANG Y J, et al. WiAct: a passive Wi-Fi-based human activity recognition system[J]. *IEEE Sensors Journal*, 2020, 20(1): 296-305.
- [12] AKHTAR Z U A, WANG H. Wi-Fi-based driver's activity recognition using multi-layer classification[J]. *Neurocomputing*, 2020(405): 12-25.
- [13] ZOU H, ZHOU Y X, YANG J F, et al. DeepSense: device-free human activity recognition via auto encoder long-term recurrent convolutional network[C]//*Proceedings of 2018 IEEE International Conference on Communications (ICC)*. Piscataway: IEEE Press, 2018: 1-6.
- [14] ZHOU Z Y, LIU C, YU X D, et al. Deep-WiID: Wi-Fi-based contactless human identification via deep learning[C]//*Proceedings of 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*. Piscataway: IEEE Press, 2020: 877-884.
- [15] SHEN X, NI Z, LIU L, et al. WiPass: 1D-CNN-based smartphone keystroke recognition Using Wi-Fi signals[J]. *Pervasive and Mobile Computing*, 2021(73): 101393.
- [16] JIA L Y, GU Y, CHENG K, et al. BeAware: convolutional neural network(CNN) based user behavior understanding through Wi-Fi channel state information[J]. *Neurocomputing*, 2020, 397: 457-463.
- [17] ZENG Y W, WU D, XIONG J, et al. Boosting Wi-Fi sensing performance via CSI ratio[J]. *IEEE Pervasive Computing*, 2021, 20(1): 62-70.
- [18] 郝占军, 段渝, 党小超, 等. 基于信道状态信息的人体复杂动作识别方法[J]. *计算机工程*, 2020, 46(1): 286-293.
HAO Z J, DUAN Y, DANG X C, et al. Human complex motion recognition method based on channel state information[J]. *Computer Engineering*, 2020, 46(1): 286-293.
- [19] 郑增威, 杜俊杰, 霍梅梅, 等. 基于可穿戴传感器的人体活动识别研究综述[J]. *计算机应用*, 2018, 38(5): 1223-1229, 1238.
ZHENG Z W, DU J J, HUO M M, et al. Review of human activity recognition based on wearable sensors[J]. *Journal of Computer Applications*, 2018, 38(5): 1223-1229, 1238.
- [20] YATANI K, TRUONG K N. BodyScope: a wearable acoustic sensor for activity recognition[C]//*Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. New York: ACM Press, 2012: 341-350.
- [21] CHEN Z H, ZHU Q C, SOH Y C, et al. Robust human activity recognition using smartphone sensors via CT-PCA and online SVM[J]. *IEEE Transactions on Industrial Informatics*, 2017, 13(6): 3070-3080.
- [22] SADANAND S, CORSO J J. Action bank: a high-level representation of activity in video[C]//*Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE Press, 2012: 1234-1241.
- [23] CHAKRABORTY I, ELGAMMAL A, BURD R S. Video based activity recognition in trauma resuscitation[C]//*Proceedings of 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. Piscataway: IEEE Press, 2013: 1-8.
- [24] CHEN Y, YU L, OTA K, et al. Hierarchical posture representation for robust action recognition[J]. *IEEE Transactions on Computational Social Systems*, 2019, 6(5): 1115-1125.
- [25] SIGG S, BLANKE U, TRÖSTER G. The telepathic phone: frictionless activity recognition from Wi-Fi-RSSI[C]//*Proceedings of 2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. Piscataway: IEEE Press, 2014: 148-155.
- [26] GU Y, REN F J, LI J. PAWS: passive human activity recognition based on Wi-Fi ambient signals[J]. *IEEE Internet of Things Journal*, 2016, 3(5): 796-805.
- [27] HALPERIN D, HU W J, SHETH A, et al. Tool release: gathering 802.11n traces with channel state information[J]. *ACM SIGCOMM Computer Communication Review*, 2011, 41(1): 53.
- [28] SEN S, LEE J, KIM K H, et al. Avoiding multipath to revive inbuilding Wi-Fi localization[C]//*Proceedings of MobiSys'13: Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. New York: ACM Press, 2013: 249-262.
- [29] ZHANG J, WU F X, WEI B, et al. Data augmentation and dense-LSTM for human activity recognition using Wi-Fi signal[J]. *IEEE Internet of Things Journal*, 2021, 8(6): 4628-4641.
- [30] 郭浩雨, 冯秀芳. 基于 Bi-LSTM 的 CSI 手势识别算法[J]. *计算机工程与设计*, 2022, 43(9): 2614-2621.
GUO H Y, FENG X F. CSI gesture recognition algorithm based on Bi-LSTM[J]. *Computer Engineering and Design*, 2022, 43(9): 2614-2621.
- [31] HASEGAWA R, UCHIYAMA A, HIGASHINO T. Maneuver classification in wheelchair basketball using inertial sensors[C]//*Proceedings of 2019 Twelfth International Conference on Mobile Computing and Ubiquitous Network (ICMU)*. Piscataway: IEEE Press, 2020: 1-6.
- [32] YAO B Q, GAO H, SU X. Human motion recognition by three-view kinect sensors in virtual basketball training[C]//*Proceedings of 2020 IEEE REGION 10 CONFERENCE (TENCON)*. Piscataway: IEEE Press, 2020: 1260-1265.
- [33] 熊小樵, 冯秀芳, 丁一. 基于 CSI 的手势识别方法研究[J]. *计算机应用与软件*, 2022, 39(1): 181-187.
XIONG X Q, FENG X F, DING Y. Research on hand gesture recognition method based on CSI[J]. *Computer Applications and Software*, 2022, 39(1): 181-187.
- [34] LIU W, CHANG S, LIU Y, et al. Wi-PSG: detecting rhythmic movement disorder using COTS Wi-Fi[J]. *IEEE Internet of Things Journal*, 2021, 8(6): 4681-4696.
- [35] WAYT H J, KHAN T R. Integrated savitzky-golay filter from inverse taylor series approach[C]//*International Conference on Digital Signal Processing*. Piscataway: IEEE Press, 2007: 375-378.

- [36] QIAN K, WU C S, YANG Z, et al. PADS: Passive detection of moving targets with dynamic speed using PHY layer information[C]//Proceedings of 2014 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS). Piscataway: IEEE Press, 2015: 1-8.
- [37] KRESGE K, MARTINO S, ZHAO T M, et al. Wi-Fi-based contactless gesture recognition using lightweight CNN[C]//Proceedings of 2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS). Piscataway: IEEE Press, 2021: 645-650.
- [38] DENIL M, BAZZANI L, LAROCHELLE H, et al. Learning where to attend with deep architectures for image tracking[J]. Neural Computation, 2012, 24(8): 2151-2184.
- [39] GUO L L, WANG L, LIU J L, et al. A novel benchmark on human activity recognition using Wi-Fi signals[C]//Proceedings of 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom). Piscataway: IEEE Press, 2017: 1-6.
- [40] LIU S Q, ZHAO Y C, CHEN B. WiCount: a deep learning approach for crowd counting using Wi-Fi signals[C]//Proceedings of 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC). Piscataway: IEEE Press, 2018: 967-974.
- [41] CHEN Z H, ZHANG L, JIANG C Y, et al. Wi-Fi CSI based passive human activity recognition using attention based BLSTM[J]. IEEE Transactions on Mobile Computing, 2019, 18(11): 2714-2724.

[作者简介]



刘洋（1996-），男，齐鲁工业大学（山东省科学院）计算机科学与技术学院硕士生，主要研究方向为深度学习、Wi-Fi 感知等。



董安明（1982-），男，博士，齐鲁工业大学（山东省科学院）副教授，主要研究方向为通信信号处理、MIMO 无线通信、机器学习、智能物联网等。



禹继国（1972-），男，博士，齐鲁工业大学（山东省科学院）教授，主要研究方向为智能感知、无线网络与通信、网络与数据安全及隐私保护、区块链、分布式计算等。



赵恺（1984-），男，博士，中国科学院自动化研究所副研究员，主要研究方向为深度学习、决策智能、自主作业机器人等。



周酉（1988-），男，博士，山东海看新媒体研究院有限公司高级工程师，主要研究方向为多媒体智能信息处理、机器学习、大数据分析等。