

WiLCOUNT:一种适用于无线感知场景的轻量级人数识别模型

段鹏松, 张伊航, 方焘, 曹仰杰, 王超

引用本文

段鹏松, 张伊航, 方焘, 曹仰杰, 王超. [WiLCOUNT:一种适用于无线感知场景的轻量级人数识别模型](#)[J]. 计算机科学, 2025, 52(10): 317-327.

DUAN Pengsong, ZHANG Yihang, FANG Tao, CAO Yangjie, WANG Chao. [WiLCOUNT:A Lightweight Crowd Counting Model for Wireless Perception Scenarios](#) [J]. Computer Science, 2025, 52(10): 317-327.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多尺度深度可分离ResNet的废弃家电回收图像分类模型](#)

Image Classification Model for Waste Household Appliance Recycling Based on Multi-scaleDepthwise Separable ResNet

计算机科学, 2025, 52(6A): 240500057-7. <https://doi.org/10.11896/jsjcx.240500057>

[基于可分离卷积与小波变换融合的道路裂缝检测](#)

Road Crack Detection Based on Separable Convolution and Wave Transform Fusion

计算机科学, 2024, 51(11A): 240100141-9. <https://doi.org/10.11896/jsjcx.240100141>

[零边界条件下一维非线性细胞自动机可逆性的判定算法](#)

Decision Algorithms for Reversibility of One-dimensional Non-linear Cellular Automata Under Null Boundary Conditions

计算机科学, 2024, 51(10): 330-336. <https://doi.org/10.11896/jsjcx.240100207>

[面向电台通信的CLU-Net语音增强网络](#)

CLU-Net Speech Enhancement Network for Radio Communication

计算机科学, 2024, 51(9): 338-345. <https://doi.org/10.11896/jsjcx.230700200>

[事件抽取方法综述:深度学习与预训练对比分析](#)

Survey on Event Extraction Methods:Comparative Analysis of Deep Learning and Pre-training

计算机科学, 2024, 51(9): 196-206. <https://doi.org/10.11896/jsjcx.231000123>

WiLCount:一种适用于无线感知场景的轻量级人数识别模型

段鹏松 张伊航 方焘 曹仰杰 王超

郑州大学网络空间安全学院 郑州 450003

(duanps@zzu.edu.cn)

摘要 针对CSI中空间特征缺失导致人数识别模型精度有限且计算复杂度较高的问题,提出了一种基于幅相融合的轻量级人数识别模型WiLCount。首先,针对原始相位信息中存在载波频率偏移和采样频率偏移而无法直接使用的问题,使用线性变换方法对相位信息进行校准;其次,将幅相数据重构为二维图像,以充分利用CSI信息中蕴含的人数空间映射特征;最后,融合深度可分离卷积与多分支结构技术,设计了一种轻量级的人数识别模型WiLCount。目前,在Wi-Fi感知人数领域暂无公开数据集,为此精心构建了一个在人数规模、行为种类均处于业界领先水平的自采数据集,并已公开。实验结果表明,WiLCount在自采数据集上的识别准确率高达99.58%,参数规模仅为同类模型的4%,相比现有方法有显著提升,且具有较好的鲁棒性。

关键词: Wi-Fi感知;信道状态信息;人数识别;幅相融合;深度可分离卷积

中图分类号 TP393

WiLCount: A Lightweight Crowd Counting Model for Wireless Perception Scenarios

DUAN Pengsong, ZHANG Yihang, FANG Tao, CAO Yangjie and WANG Chao

School of Cyber Science and Engineering, Zhengzhou University, Zhengzhou 450003, China

Abstract To address the challenges of limited accuracy and high computational complexity in crowd counting models due to the absence of spatial features in CSI, this paper proposes a lightweight model, WiLCount, based on amplitude-phase fusion. Firstly, a linear transformation method is applied to calibrate the phase data, addressing the issues of carrier frequency offset and sampling frequency offset in the raw phase information, which would otherwise render it unusable. Next, the amplitude-phase data is reconstructed into a two-dimensional image to fully exploit the spatial mapping features of crowd count inherent in CSI data. Finally, WiLCount is developed by integrating depthwise separable convolutions with a multi-branch structure. Due to the lack of publicly available datasets in the Wi-Fi-based crowd counting field, a self-collected dataset, leading the industry in terms of crowd scale and activity diversity, is meticulously constructed and released. Experimental results demonstrate that WiLCount achieves a recognition accuracy of up to 99.58% on the self-collected dataset, with a parameter size of only 4% of that of comparable models. Significant improvements over existing methods have been observed, with the model exhibiting strong robustness.

Keywords Wi-Fi sensing, Channel state information, Crowd counting, Amplitude-phase fusion, Depthwise separable convolution

1 引言

人数识别作为人机交互的前提,是人工智能领域的重要研究方向之一,一直深受研究人员的关注。人数识别指通过对目标环境的监测,对场景中的个体数量进行统计或估计。随着人工智能技术的快速发展,人数识别在多个领域展示出了巨大的应用潜力。

过去,较常用的人数识别方法有基于专用传感器(如雷达^[1]、红外传感器^[2])和基于计算机视觉^[3-4]两类。其中,基于专用传感器的方法通常能够提供高频感知数据,适合对实时

性要求较高的应用场景,但携带不便且部署成本较高。基于计算机视觉的方法通过提取图像或视频数据特征,可以实现更精准的人数识别,但易受光线遮挡影响且隐私保护性较差。这些缺陷限制了人数识别的应用范围,在特殊场景下尤为明显。例如,在日常居家场景中,需要对人们的隐私进行保护;在政府机关、大型商场等重要场所中,存在夜间的弱光照等问题。针对以上受限场景,传统的人数识别方法显然难以满足。

近年来,随着Wi-Fi设备的大规模应用,为使用Wi-Fi信号感知周围场景提供了较好的感知环境。凭借低成本、隐私性好、部署广泛、不受光照影响等优势,基于Wi-Fi信号的

到稿日期:2024-08-12 返修日期:2024-10-15

基金项目:郑州市协同创新重大专项(20XTZX06013);河南省自然科学基金(222300420295);中国工程科技发展战略河南研究院战略咨询研究项目(2022HENYB03);河南省科技攻关项目(232102210050)

This work was supported by the Zhengzhou Collaborative Innovation Major Project(20XTZX06013), Natural Science Foundation of Henan Province(222300420295), China Engineering Science and Technology Development Strategy Henan Research Institute Strategic Consulting Research Project(2022HENYB03) and Henan Province Science and Technology Research Project(232102210050).

通信作者:曹仰杰(caoyj@zzu.edu.cn)

无线感知技术,为智能人机交互应用提供了一种新的解决思路,受到研究者的青睐^[5]。目前,此类无线感知研究已初步应用于手势识别^[6]、行为识别^[7]、步态识别^[8]、呼吸检测^[9]、目标跟踪^[10]等不同领域。

与此同时,基于 Wi-Fi 信号的人数识别相关研究也在持续开展中。Bahl 等^[11]提出基于 Wi-Fi 接收信号强度指示(Received Signal Strength Indicator, RSSI)的室内定位系统 Radar,激发了研究人员利用 RSSI 来解决人数识别问题^[12-16]的兴趣。然而, RSSI 是 MAC 层信号,易受室内环境和多径效应影响,导致识别结果存在较大的不稳定性。2011 年, Halperirr 等^[17]开源了 CSI Tool 工具,为研究人员从商用 Wi-Fi 设备中提取 CSI 提供了极大便利。信道状态信息(Channel State Information, CSI)是对物理层信号的细粒度测量,它描述了信道中每个正交子载波的振幅和相位。Yang 等^[18]通过对比分析 RSSI 和 CSI,证明了 CSI 可以区分多径效应。因此,相较于 RSSI, CSI 更适用于人数识别,且稳定性更好。目前,已有不少学者利用 CSI 信息进行人数识别^[19-25]。

然而,充分利用 CSI 并不容易。虽然 CSI 同时携带信道的振幅和相位信息,但现有相关人数识别方法大多数仅利用振幅信息,而相位信息往往被忽视^[19-22],存在特征缺失的风险。从原始 CSI 数据中提取到的相位信息中存在载波频率偏移(Carrier Frequency Offset, CFO)和采样频率偏移(Sampling Frequency Offset, SFO),导致相位信息难以直接使用。目前基于 CSI 的人数识别研究中,为了简化数据处理流程,大多数将采集到的 CSI 作为一维时序数据流直接使用,未能充

分利用不同子载波间的空间特征,使得高精度人数识别难以实现。此外,人数识别通常应用于边缘侧,算法或模型运行载体存在资源受限的约束。为此,本文设计了一种轻量级非接触式人数识别模型,用于解决上述问题。本文主要贡献如下:

1)针对相位信息难以利用且载波间空间特征缺失的问题,通过线性变换对相位信息进行校准,并提出一种幅相特征重构机制。该机制将一维时序 CSI 数据重构为二维振幅和相位图像数据,使其同时包含子载波内和子载波间的人数特征,为后续识别过程提供了丰富的特征信息。

2)针对现有方法计算复杂度较高的问题,提出一种基于深度学习的轻量化人数识别模型 WiLCount。该模型采用幅相特征重构机制,并构建深度可分离卷积与多分支结构组合模块,通过逐通道卷积提取特征。实验结果表明, WiLCount 在自采数据集上的识别准确率高达 99.58%,且模型参数量仅为同类模型的 4%,具有显著优势。

本文数据集和主要代码已上传 GitHub¹⁾。

2 相关工作

基于 Wi-Fi 信号的人数识别可根据所使用的数据类型不同划分为两类:基于 RSSI 的人数识别和基于 CSI 的人数识别。在基于 RSSI 的方法中,主要依赖测量 Wi-Fi 信号的强度变化,其通过分析信号强度的波动来推断人数的变化。相比之下,基于 CSI 的人数识别则对更为精细的信道状态信息(含振幅和相位)变化进行分析,从而识别人数。该部分的主要相关工作如表 1 所列。

表 1 相关工作
Table 1 Related works

文献(发表年份)	使用数据	方法	性能
[12](2008)	RSSI 均值、方差	线性回归	20 人;平均计数误差 1.5 人
[13](2015)	RSSI	支持向量回归	7 人;77.2%
[14](2015)	RSSI	概率分布模型	9 人;计数误差 2 人及以下情况占 63%
[15](2015)	RSSI 均值、方差	神经网络	25~50 人;80%
[16](2018)	RSSI	决策树	4 人;98.2%
[19](2014)	CSI PEM	Grey Verhulst 模型	30 人;计数误差 2 人及以下情况占 98%
[20](2019)	CSI 振幅	KNN	4 人;90%
[21](2022)	CSI 振幅	LSTM	5 人;97%
[22](2022)	CSI 振幅	机器学习	小型会议室 5 人;计数误差 1 人情况占 89.9% 中型会议室 10 人;计数误差 1 人情况占 81.8%
[23](2022)	CSI 相位差	BP 神经网络	5 人;90%以上
[24](2022)	CSI 振幅、相位差	小样本学习	8 人;92.68%
[25](2023)	CSI 振幅	迁移学习	5 人;82%

2.1 基于 RSSI 的人数识别

当人体在从发射端到接收端的视距路径中移动时, RSSI 会出现波动^[26]。随着感知范围内待识别人数的增加,对无线信号的扰动影响也逐渐增大。人数变化和 RSSI 信号波动间存在的非线性关系,使得利用 RSSI 进行人数识别成为可能^[12]。Nakatuka 等^[12]通过线性回归提出了一种使用 RSSI 的人数识别系统,其平均计数误差为 1.5 人。Yoshida 等^[13]通过分析 RSSI 数据,提出一种基于支持向量回归的人数识别方法,在室内不超过 7 人的情况下,准确率达到了 77.2%。Depratla 等^[14]在限定区域内部署一对 Wi-Fi 收发天线,以 RS-

SI 测量值为系统输入,提出一种概率分布模型对区域内人数进行识别。实验结果表明,在室内使用全向天线进行统计时,该方法在 63%的情况下误差小于 2 人。Li 等^[15]提出了 Wi-Counter,其根据路由器与检测地点的视距以及非视距分类来计算 RSSI 的均值与方差以检测人员的数量。Kura 等^[16]使用决策树算法对 RSSI 数据进行分类,得到走廊场景人数的估计值。

然而,由于 RSSI 是一种粗粒度的信息,极易受到室内环境中多径效应的影响,且随着室内人数的增加,干扰会进一步加剧。在复杂的室内环境中, RSSI 无法提供足够的区分度,

¹⁾ <https://github.com/zzyZYH/WiLCount>

导致使用 RSSI 的人数识别算法的准确率会急剧下降。

2.2 基于 CSI 的人数识别

与 RSSI 相比,CSI 可提供多个子载波的细粒度信息,在相同室内环境中具有更高的稳定性。人数识别研究中,可以提取由人数变化而引起的 CSI 测量值扰动,通过对 CSI 进行分析进而实现人数预测。

Xi 等^[19]提出了一种基于 CSI 的人数识别方法。该方法通过扩展 CSI 矩阵中非零元素的百分比(Percentage Of Non-zero Element, PEM)来表示无线信道的变化,并通过 Grey Verhulst 模型建立 PEM 与人数之间的单调关系。在误差不超过 2 个人的情况下,该方法在室内和室外的识别准确率分别达到 98% 和 70%。Ding 等^[20]通过 CSI 幅值波动刻画室内人数的变化,使用 KNN 分类器来识别人数。在实验室环境不超过 4 人的情况下,该方法的准确率达到 90%;在会议室环境不超过 2 人的情况下,该方法的准确率达到 89.58%。Wang 等^[21]提出了一种基于 CSI 的人数识别方法。该方法使用 LSTM 算法在 3 个真实的实验场景下进行了人数统计实验,在不超过 5 人的情况下,总体准确率超过 97%。Choi 等^[22]提出使用 ESP32 节点构建 Wi-Fi 链路进行人群计数和定位的系统 Wi-CaL。在误差范围为 1 人的情况下,Wi-CaL 在小型和中型会议室中的识别率分别为 89.8% 和 81.8%。Guo 等^[23]提出了一种使用 Wi-Fi 信号的穿墙式人群计数系统 TWCC。TWCC 提取 CSI 的相位差数据,并对其进行处理后从中提取时域、子载波域、频域和时频域特征,对不同域特征使用不同的 BP 神经网络,再结合权值调整和阈值划分,使得其整体平均识别精度约为 90%。Hou 等^[24]提出一种在新环境中使用有限数据样本获得极高跨域检测精度的框架 DASECount。实验结果表明,DASECount 利用小样本学习,在 0~8 人的计数任务中实现了超过 92.68% 的检测准确率。Khan 等^[25]利用迁移学习方法,通过将源环境中的训练模型用于其他室内环境,实现了人数识别功能。结果表明,与传统的卷积神经网络相比,该方法在缩短 40% 训练时间的同时还能提升 4.7% 的精度。

目前,大多数基于 CSI 的人数识别方法仅仅利用振幅信息,而对相位信息利用不足,影响识别效果。对于室内场景下的人数识别来说,其相位信息包含了潜在的个体间相对位置关系信息,这对识别结果具有重要的作用。

3 模型原理

3.1 信道状态信息

基于 Wi-Fi 信号的无线感知技术通常仅需部署一对发射 Wi-Fi 信号的路由器和接收信号的设备。在典型的室内环境中,由于无线信道的多径效应,发射端发出的无线信号在到达接收端时会经历多条传播路径,如图 1 所示。当发射端与接收端之间无障碍物时,信号中至少存在一条直达路径,即视距路径(Line Of Sight, LOS);其余信号则通过天花板、地板、墙壁等物体反射,形成多条非视距路径(Non-Line Of Sight, NLOS)。当人体进入该室内环境时,原有的信号传播路径会因人体的存在而发生扰动,从而改变信号的传播特性。由于

此类扰动,接收端接收到的 Wi-Fi 信号会携带反映环境变化的传输特征,进而能够用于感知目标的存在与运动。

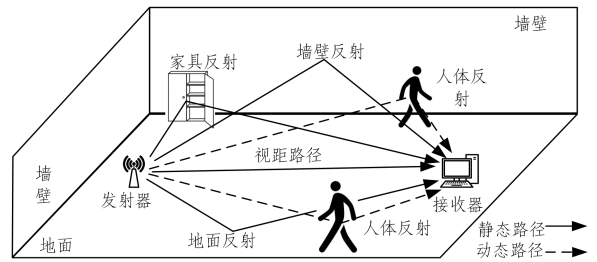


图 1 信号传播图

Fig. 1 Signal propagation diagram

CSI 描述了 Wi-Fi 信号在无线网络中不同传播路径下的变化特征。当人体在监测范围内运动时,会影响 Wi-Fi 信号传播路径上的折射、反射和散射等变化,导致接收到的 CSI 信息发生变化。接收到的 CSI 是一个数据序列,其中每个值代表当前时刻所有子载波的信道响应信息,其可定义为:

$$\mathbf{H} = [\mathbf{H}(f_k, t_1), \mathbf{H}(f_k, t_2), \dots, \mathbf{H}(f_k, t_n)] \quad (1)$$

如果采用 Halperin 等^[17]开发的 Linux CSI Tool 工具获取原始 CSI 数据,则 t 时刻的 CSI 由 IEEE 802.11n 标准中 56 个 OFDM 子载波中的 30 个子载波组成。每个子载波均有相应的包含振幅和相位信息的 CSI 值,可表示为:

$$\mathbf{H}(f_k, t_n) = \sum_{i=1}^N |\mathbf{H}_i(f_k, t)| e^{i\angle \mathbf{H}_i(f_k, t)}, k \in [1, k] \quad (2)$$

其中, N 为路径数, $|\mathbf{H}_i(f_k, t)|$ 与 $\angle \mathbf{H}_i(f_k, t)$ 分别代表第 k 个子载波在 t 时刻的振幅与相位。在人数识别研究中,人数变化对 Wi-Fi 信号的扰动具体体现为对 CSI 中相位和振幅的影响。

3.2 问题定义

在使用 CSI Tool 工具采集 CSI 数据时,获得的数据格式为 $N_s \times N_r \times n$ 的复数矩阵,其中 N_s, N_r 分别表示发射器和接收器天线数量, n 表示子载波数量。为了对感知环境中的人数进行识别,需要采集时长 T 的 CSI 数据,从而得到一个维度为 $N_s \times N_r \times n \times T$ 的矩阵。为更直观地定义问题,可将式(2)中的 CSI 数据直观表示为:

$$\mathbf{CSI} = (\mathbf{CSI}_{\text{amplitude}}, \mathbf{CSI}_{\text{phase}}) \quad (3)$$

其中, $\mathbf{CSI}_{\text{amplitude}}$ 为 CSI 的振幅矩阵,包含 30 个子载波的振幅信息; $\mathbf{CSI}_{\text{phase}}$ 为 CSI 的相位矩阵,也包含 30 个子载波的相位信息。此时,幅相融合问题可抽象为求振幅信息 $\mathbf{CSI}_{\text{amplitude}}$ 和相位信息 $\mathbf{CSI}_{\text{phase}}$ 的融合方法,可形式化为:

$$\mathbf{X} = \text{fusion}(\mathbf{CSI}_{\text{amplitude}}, \mathbf{CSI}_{\text{phase}}) \quad (4)$$

对于人数识别,可将包含人数信息的数据集表示为 $D = (X_i, Y_i), i \in [1, I]$,其中 D 中共包含 I 条 CSI 数据, X_i 和 Y_i 分别表示融合后的振幅与相位特征及其对应标签。在数据集构建完成后,通过提取融合特征中的有效信息并进行分类,可得到高精度的人数预测结果。

4 模型设计

以重构后的振幅和相位数据为输入,本文设计了一个基于深度学习的轻量级人数识别模型 WiLCount,其总体流程如图 2 所示。

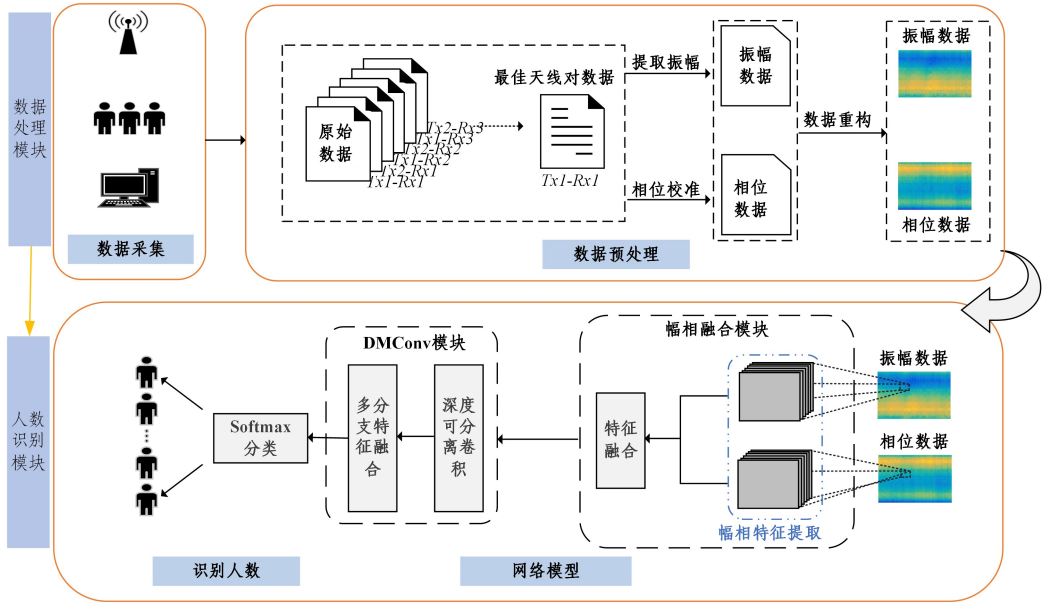


图2 WiLCount的总体流程

Fig. 2 Overall flow chart of WiLCount

WiLCount 主要由数据处理模块和人数识别模块组成。在数据处理模块中,主要进行 CSI 数据采集与预处理,包括数据采集、最佳天线对选择、相位校准以及数据重构 4 个步骤。人数识别模块则负责将经过处理的 CSI 数据输入设计好的深度神经网络 WiLCount 中。该模块中,系统从幅相数据中提取出关键的人数特征,并对这些特征进行分类,以实现对不同人数的准确识别。数据采集部分的相关内容在 5.1.1 节中进行介绍,本章将主要聚焦于数据处理和 WiLCount 人数识别模型的设计细节。

4.1 数据处理

4.1.1 最佳天线对选择

本文实验中,在发射端与接收端分别使用 2 根和 3 根可检测天线。每根发射天线与接收天线之间形成一条数据流,因此共有 $6(2 \times 3)$ 条 CSI 数据流。6 条数据流之间存在较大差异,因此在进行数据选择时,需要确定最佳天线对。本文中,通过选择最大信噪比(Signal to Noise Ratio, SNR)的方法得到最佳天线对^[27]。第 i 个天线对的 SNR 表示为:

$$\text{SNR}(\text{ANT}_s) = \max\{10 \times \log_{10}(P_{S_i}/P_{N_i})\}, i = 1, \dots, 6 \quad (5)$$

其中, P_{S_i} 和 P_{N_i} 是一个天线对中 30 个子载波的信号和噪声的平均功率。

4.1.2 相位校准

尽管 CSI 已经在各种研究中得到广泛应用,但很多研究主要关注于利用 CSI 的振幅信息,而对 CSI 的相位信息往往并不重视。主要原因是,原始 CSI 数据中提取的相位信息受到了载波频率偏移和采样频率偏移的影响,使其难以直接应用^[28]。为了充分利用相位信息,本文采用线性变换^[29]的方法对相位信息进行校准。

在 CSI 中,第 i 个子载波上的原始相位 $\hat{\phi}_i$ 可表示为:

$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N} \delta + \beta + Z \quad (6)$$

其中, ϕ_i 为真实相位, δ 表示由 SFO 引起的时序偏移, β 表

示由 CFO 引起的未知相位偏移, Z 为测量噪声, k_i 是第 i 个子载波的索引, N 是快速傅里叶变换的长度。为减小相位误差造成的影响,对原始相位进行线性变换操作,通过从原始相位 $\hat{\phi}_i$ 中减去 $ak_i + b$, 来消除 δ 和 β 造成的误差,得到校准后的相位 $\tilde{\phi}_i$ 。其中 $a = \frac{\hat{\phi}_n - \hat{\phi}_1}{k_n - k_1} = \frac{\phi_n - \phi_1}{k_n - k_1} - \frac{2\pi}{N} \delta$ 为接收到的原始相位斜率; $b = \frac{1}{n} \sum_{j=1}^n \hat{\phi}_j + \beta$ 为偏移量。此时,校准后的相位 $\tilde{\phi}_i$ 表示为:

$$\tilde{\phi}_i = \hat{\phi}_i - ak_i - b = \phi_i - \frac{\phi_n - \phi_1}{k_n - k_1} k_i - \frac{1}{n} \sum_{j=1}^n \hat{\phi}_j \quad (7)$$

相位校准前后的数据对比如图 3 所示。可以看出,与原始相位相比,校准后的相位分布相对稳定,可作为人数识别中的参考特征。

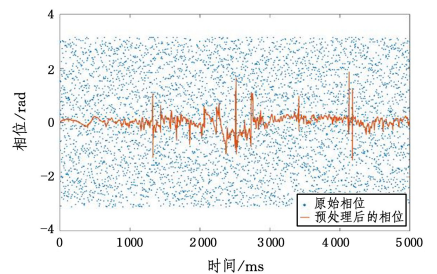


图3 相位校准

Fig. 3 Phase correction

4.1.3 数据重构

对相位信息进行校准后,可从原始 CSI 数据中获取准确有效的振幅和相位信息。但值得注意的是,CSI 子载波对人体的敏感度是不一致的,单个子载波的敏感度都呈现不稳定的特性^[30]。因此,每个子载波都蕴含着独特的人数感知特征。为充分利用子载波间 CSI 特性的差异,将振幅和相位信息分别重构为二维矩阵,进而从中提取出更多空间特征。其中振幅信息是从原始 CSI 数据中直接提取的,相位信息是从

原始 CSI 数据中提取后通过消除相位误差得来的。该二维矩阵的尺寸为 $N_t \times N_c$, 其中 N_t 为采样点个数, N_c 为子载波

个数。对二维矩阵使用 jet 方案进行着色后,可以实现可视化呈现,如图 4 所示。

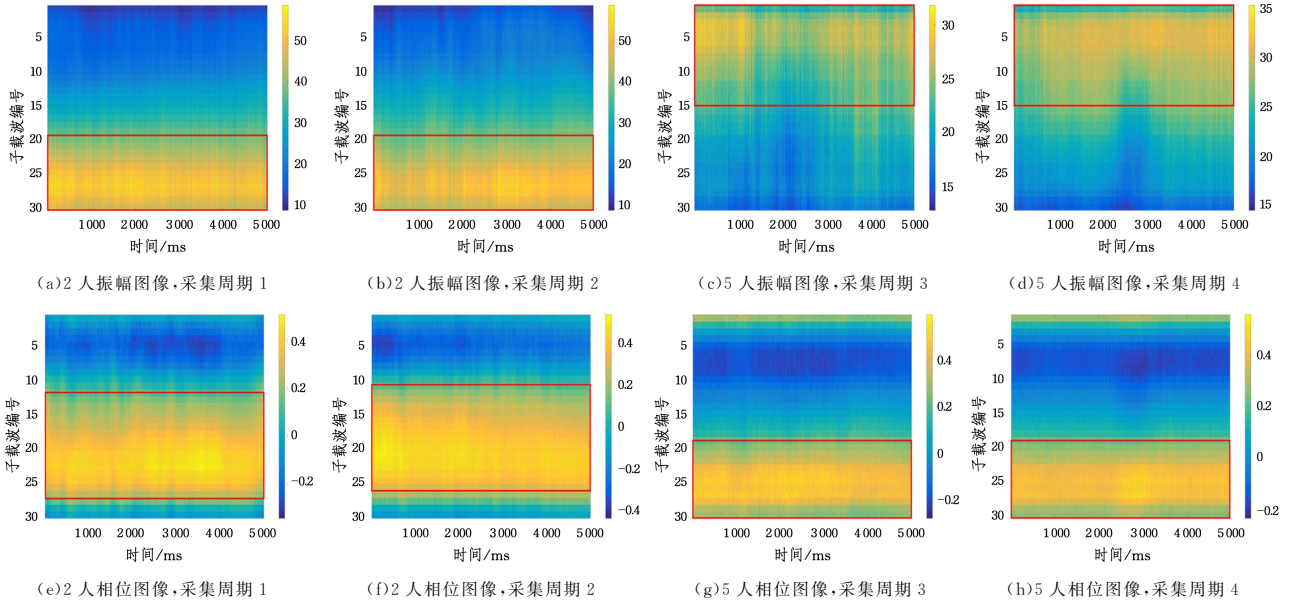


图 4 不同采集周期内不同人数振幅和相位图像

Fig. 4 Amplitude and phase images of different numbers of people in different acquisition cycles

通过重构后的二维图像数据可以观察到,相同人数的感知数据呈现出相似特征,而不同人数的感知数据则有明显差异。原因在于,环境中的非人体障碍物和人体对 Wi-Fi 信号的传播影响不同。在相同人数的情况下,环境和障碍物的分布相对固定,使得信号的传播路径相对稳定,因此二维图像呈现出相似特征。在不同人数的情况下,人体对信号的吸收和反射会导致信号传播路径变化,从而使得二维图像呈现出明显差异。因此,重构后的二维图像数据能够有效

反映 Wi-Fi 信道中所携带的时频特征与空间特征。模型能够有效地从振幅特征和相位特征中提取更为丰富的时频域信息与空间域信息。通过使用这些联合特征,能够更准确地识别人数。

4.2 人数识别

以重构后的振幅和相位数据为输入,本文设计了一种幅相融合轻量化模型 WiLCount。WiLCount 主要由幅相融合子模块与 DMConv 子模块构成,其结构如图 5 所示。

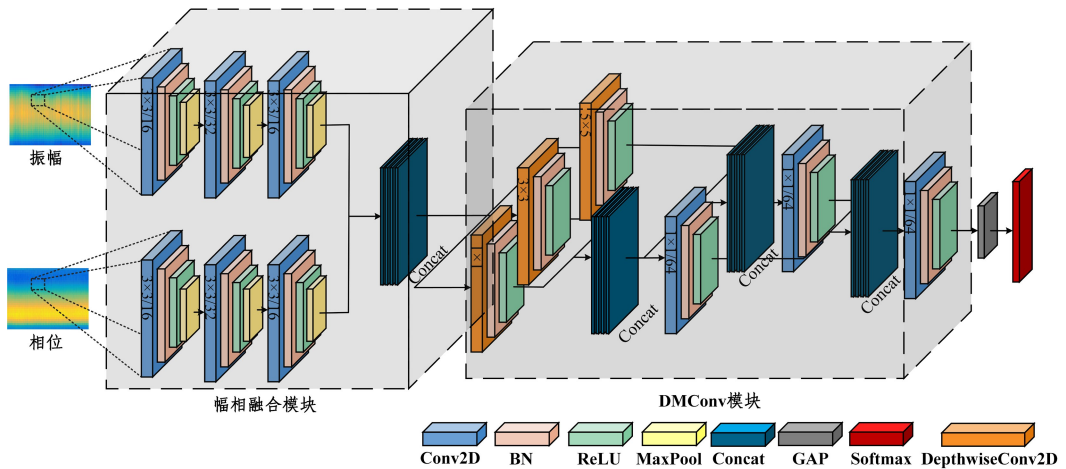


图 5 WiLCount 结构

Fig. 5 Structure of WiLCount

4.2.1 幅相融合子模块

幅相融合子模块如图 6 所示。首先,使用两个独立分支分别用于提取振幅与相位的初级特征。这两个分支结构相同,由 3 个串联卷积模块构成。每个卷积模块由卷积层、归一化层、线性整流层和最大池化层 4 部分组成。3 个卷积模块均使用大小为 3×3 的卷积核,个数分别为 16, 32 和 64。该设

计的目的是通过多层次的卷积操作,使模型能够逐步提取和组合更为细致的特征。其次,使用归一化层加快网络收敛速度,随后使用线性整流层增强网络的非线性能力,以更好地拟合训练数据。最后,加入最大池化层去除冗余信息,对特征进行压缩。通过两个独立分支从振幅与相位信息中提取出初级特征,再对这两个分支的输出进行合并,以更好地捕捉感知

数据的关键特征,提高模型识别性能。

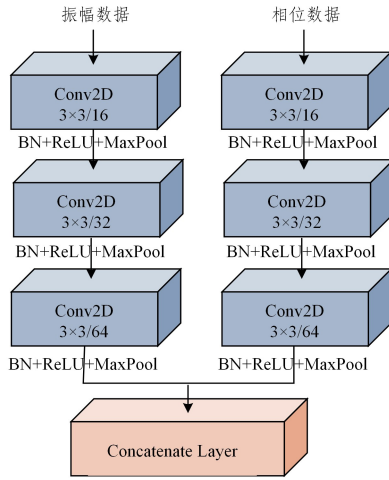


图6 幅相融合子模块结构

Fig. 6 Structure of amplitude-phase fusion submodule

4.2.2 DMConv 子模块

DMConv 子模块如图 7 所示,DMConv 子模块中采用深度可分离卷积^[31]与多分支结构的设计。在使用深度可分离卷积时,对于每个输入通道,首先使用单独的卷积核对其进行卷积操作,这使得每个通道只需要学习与其相关的特征,而不需要关注其他通道的信息。其次,使用 1×1 的卷积核在不同通道之间进行卷积,用于将通道之间的信息进行混合。在整个过程中,卷积核的参数得以共享,每个卷积核只学习部分特征。相比之下,传统卷积需要每个卷积核学习所有通道的特征,导致参数数量大幅增加。本文设计策略不仅保证了模型的轻量化,同时通过各个分支采用多尺寸卷积核的策略,实现了对特征图上不同感知范围特征的获取。通过将各分支的中间特征与其他分支的高层特征进行融合,实现对图像特征的全面精细提取。

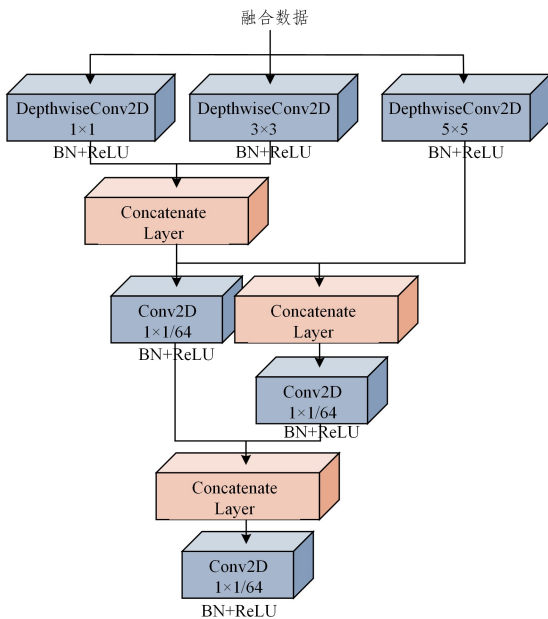


图7 DMConv 子模块的结构

Fig. 7 Structure of DMConv submodule

DMConv 子模块由 3 个分支构成,每个分支由深度可分离卷积层和归一化层、线性整流层组成,其卷积核大小分别为 1×1 , 3×3 和 5×5 。在该模块中,每个分支负责捕捉不同尺度的特征。在各分支特征融合后,再通过点卷积进行特征整合,对高阶特征进行降维,从而提高网络性能。DMConv 子模块通过引入不同分支之间的多尺度特征融合,有效地丰富了所提取的特征信息;同时,也有效地减轻了模型结构加深可能导致的模型性能下降问题。在降低参数规模的同时,可以提升识别准确率。

在 DMConv 子模块后,使用全局平均池化实现特征降维,减少模型参数。最后,通过 Softmax 函数计算不同类别的概率分布,实现对不同人数的预测。

5 实验与分析

5.1 实验设置

目前,已有多种工具用于提取 CSI 信息,包括 802.11n CSI Tool^[17], Atheros CSI Tool^[32] 和 Nexmon CSI 提取工具^[33]等。在人数识别领域,现有研究主要集中于新颖的模型结构,而缺乏公开数据集。研究人员需要自行采集数据集,这制约了研究的进一步发展。本文在室内环境下使用 802.11n CSI Tool 工具采集 Wi-Fi CSI 数据,获取环境中的人数信息并构建了一个开源数据集 Count-Fi。

为了进一步促进 Wi-Fi 感知领域的发展,填补 Wi-Fi 感知人数识别领域无公开数据集的空白,现将 Count-Fi 数据集上传 GitHub。

5.1.1 Count-Fi 数据集介绍

目前,国内外还没有公开的 Wi-Fi 感知人数识别数据集,因此本文进行了数据集的自采。针对感知数据采集装置,本文使用 TP_LINK AC1750 无线路由器作为 Wi-Fi 信号发射端,将搭载 Intel 5300 802.11n Wi-Fi NIC 网卡的台式计算机作为 Wi-Fi 信号接收端。通过在接收端设备上安装开源 802.11n CSI Tool 工具,完成对 CSI 信息的实时采集。发射端与接收端分别安装 2 根和 3 根天线,构成 $6(2 \times 3)$ 组天线对,每组天线对由 30 个子载波组成。本文实验中,设置接收端数据采样频率为 1kHz。在数据采集期间,为确保数据完整性,特别是考虑到 2.4GHz 频段同频干扰和丢包现象对数据的负面影响^[34],本文选择在 5GHz 频段进行数据采集。Count-Fi 主要的数据采集参数如表 2 所列。

表2 主要设备参数设置

Table 2 Main equipment parameter settings

参数名	参数值
发射器	TP-Link AC1750
接收器	Intel 5300 802.11n NIC
天线设置	2发3收
信号频率	5GHz
采样工具	CSI Tool
采样频率	1000 Hz

由于 Wi-Fi 信号在室内传播时受多径效应影响显著,数据采集时应选择真实、方便、干扰较少的实验环境。Count-Fi 将如图 8 所示的房间作为数据采集室。该房间中包括 4 张办

公桌,1台落地空调,发射端与接收端相距4m。为确保实验数据的可靠性,分别在不同活动状态下收集了CSI数据,包括静止和行走,如图9所示。在静止状态下,测试人员位于视距路径中心;而在行走状态下,测试人员以匀速垂直穿越视距路径行走。在两种活动状态下,分别逐渐增加测试人员的数量,从1人增加到6人,共形成7种情况(空房间、1~6人)。每种情况都进行了50次数据采集,每次采集5s。将两种活动的采集数据混合,形成实验数据集 Count-Fi。

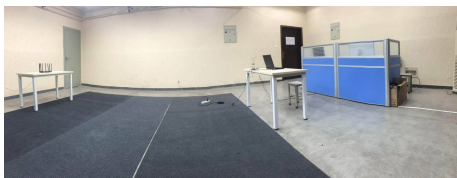
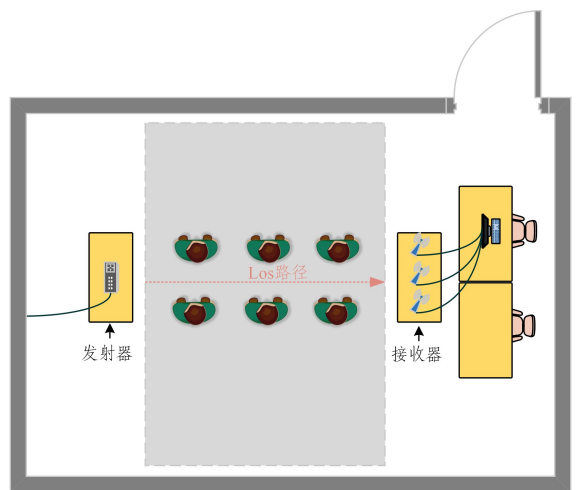
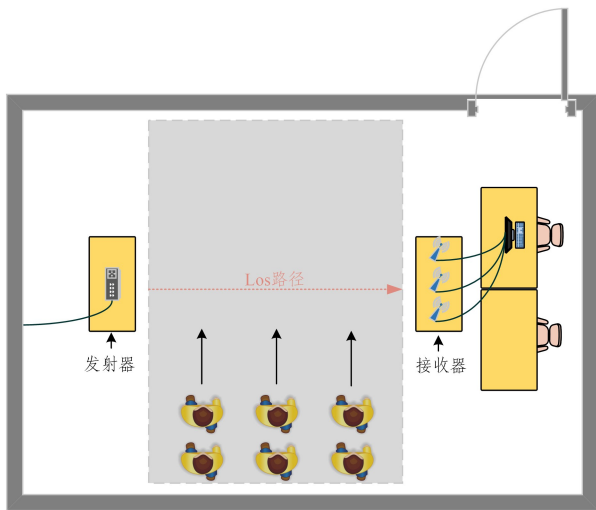


图8 采集环境

Fig. 8 Collection environment



(a) 静止



(b) 行走

图9 两种活动状态

Fig. 9 Two activity states

为扩充样本并提高模型的识别性能,采用了滑动窗口技术扩充数据集。由于滑动窗口长度与步长设置对模型的性能有显著影响,通过实验对比不同窗口长度与步长组合对模型的影响,结果如表3所列。

表3 不同滑动窗口与步长结果

Table 3 Results of different sliding windows and step sizes

滑动窗口与步长设置	准确率	精确率	召回率
1000/500	96.75	97.10	97.05
800/200	99.58	99.61	99.60
600/100	98.36	98.49	98.49

结果表明,当滑动窗口长度设置为800、步长为200时,模型的表现最佳,识别准确率达到99.58%。相较而言,当滑动窗口长度为1000、步长为500时,较大的窗口长度与步长导致模型对局部信息的捕捉能力减弱,且特征提取不足,导致识别准确率下降。当窗口长度为600、步长为100时,虽然较小的窗口和步长能够提高对细节的捕捉能力,但也引入了大量冗余数据,增加了模型的计算复杂度与训练时间,最终导致准确率的下降。滑动窗口长度为800、步长为200的组合则在确保足够特征信息的同时,避免了冗余数据的过多引入,从而有效平衡了特征提取能力与计算复杂度。

综上,使用滑动窗口长度为800ms、步长为200ms时,对采集到的数据进行分割,以扩充样本,并以8:2比例将数据划分为训练集和测试集。

为评估 Count-Fi 数据集的数据质量与规模,对近年来 Wi-Fi 感知人数识别领域的文献进行调研,调研结果如表4所列。由于现有工作中,数据集和代码未公开,导致不同工作的准确率评判标准不一,从而使得在统一基准下进行公正比较变得困难。在人数识别中,评估模型性能的方式通常有两种:精确估计和计数误差。精确估计指模型直接输出的识别人数与实际人数完全一致的准确率,而计数误差则允许一定范围内的识别偏差。例如,对于5人的场景,如果模型输出为5人或4人,则在计数误差为1的情况下被视为准确。基于计数误差的评估,Count-Fi 数据集在计数误差为1人及以下情况时,实验中的准确率达到100%。尽管现有工作中的数据集未公开,限制了与其他工作的直接对比,但通过对 Count-Fi 数据集的实验评估,本研究能够有效展示 Count-Fi 数据集在实际应用中具有较高的准确性和较为全面的数据覆盖,表明了其在数据质量、样本数量及行为类型上的显著优势,彰显出其在 Wi-Fi 感知人数识别领域处于较领先水平。

表4 与现有工作的比较结果

Table 4 Comparison results with existing work

现有工作	识别方法	识别人数	识别效果
[20]	KNN	0~4	90%
[21]	LSTM	0~5	97%
[22]	机器学习	小型会议室:0~5 中型会议室:0~10	计数误差1人情况占89.9% 计数误差1人情况占81.8%
[23]	BP神经网络	0~5	90%
[24]	小样本学习	0~8	92.68%
[25]	迁移学习	1~5	82%
WiLCount	DMConv	0~6	99.58%

5.1.2 评估指标

实验采用准确率(A)、精准率(P)和召回率(R)对识别效

果进行评估。其中,准确率表示为所有正确预测的数目与总预测数的比值;精确率表示为正确预测的正值与总预测正值的比值;召回率表示为正确预测的正值与总实际正值的比值。各指标的计算式如下:

$$Accuracy(A) = \frac{TP + TN}{TP + TN + FN + TN} \quad (8)$$

$$Precision(P) = \frac{TP}{TP + FP} \quad (9)$$

$$Recall(R) = \frac{TP}{TP + FN} \quad (10)$$

其中, TP 表示属于该类且被分为该类的样本数; TN 表示不属于该类且不被分为该类的样本数; FP 表示不属于该类但被分为该类的样本数; FN 表示属于该类但不被分为该类的样本数。

5.2 实验结果

5.2.1 WiLCount 人数识别性能

本文中,为验证 WiLCount 模型在不同人数场景下的识别性能,评估其对人员密度变化时的适应性,分别使用 1~6 人的数据进行实验,实验结果如表 5 所列。可以看出,在 1~3 人的低密度人数场景下,由于人数较少,模型的识别任务相对简单,WiLCount 模型的准确率达到 99.73% 以上。在 4~6 人的高密度人数场景下,人员密度增加导致环境信息复杂,增大了识别难度,导致其识别准确率略微下降。然而,即便在高密度场景下,准确率仍然维持在 99.39% 以上,表明 WiLCount 模型在处理多人场景时依然保持出色表现。实验结果表明,WiLCount 模型对变化的人数的密度具有强大的适应性,能够在各种密度变化下保持高准确率,并且在不同人数场景下具有较好的鲁棒性。

表 5 不同人数的实验结果

Table 5 Experimental results with different numbers of people (%)

人数	准确率	精确率	召回率
1 人	100.00	100.00	100.00
2 人	99.73	99.78	99.78
3 人	99.93	99.94	99.94
4 人	99.39	99.45	99.45
5 人	99.71	99.73	99.73
6 人	99.58	99.61	99.60

5.2.2 不同特征数据对性能的影响

本文中,WiLCount 模型结合 CSI 振幅和相位信息中的特征,设计幅相融合模块。为验证使用不同数据的人数识别性能,分别使用振幅数据、相位数据和幅相融合数据进行实验,识别性能如表 6 所列。当只使用振幅或相位数据进行实验时,将幅相融合模块中的双独立分支修改为单分支,其余结构保持不变。

表 6 不同特征数据的实验结果

Table 6 Experimental results with different feature datas (%)

数据集	准确率	精确率	召回率
振幅	98.88	98.95	98.92
相位	97.65	97.82	97.82
幅相融合	99.58	99.61	99.60

能够发现,在仅使用振幅数据的情况下,识别准确率为 98.88%。仅使用相位数据时,识别准确率为 97.65%。由于相位信息容易受到环境噪声的影响,使用振幅数据对比相位数据更容易提取可用特征。然而,当同时使用幅相融合数据时,WiLCount 的人数识别性能显著提升,准确率达到 99.58%。准确率的提升并不仅仅因为简单的数据堆叠,还有将提取的振幅和相位特征进行融合后,可以充分利用来自两种信息源的互补性。振幅主要反映信号的强度变化,对环境中的大范围扰动较为敏感,而相位则能够提供信号传播路径上的微小相位差异信息,反映更为细致的扰动变化。然而,单独使用振幅或相位信息会丢失部分信息,导致模型的特征提取能力受限。因此,通过在模型中进行幅相融合,能够有效整合两种特征,既提升了模型对宏观与微观信息的捕捉能力,又能在特征维度上实现互补,有助于模型更准确地捕捉 CSI 信息中的人数特征,从而提高整体性能。

5.2.3 模型鲁棒性测试

1) 不同活动模式对比

为验证不同活动模式对 WiLCount 模型性能的影响,分别使用静止状态数据、行走状态数据和包含两种数据的 Count-Fi 数据集进行实验对比,结果如表 7 所列。

表 7 不同活动的实验结果

Table 7 Experimental results with different activities (%)

活动模式	准确率	精确率	召回率
静止	99.98	99.98	99.98
行走	99.02	99.01	99.02
静止+行走	99.58	99.61	99.60

可以看出,当使用静止状态的数据进行实验时,WiLCount 模型的识别精度达到了 99.98%,这表明在静态的简单环境中,WiLCount 能够高效地进行人数识别。相比之下,当使用行走状态的数据时,行走状态增加了检测环境的复杂性,使得识别精度略有下降,为 99.02%。然而,当使用二者混合的 Count-Fi 数据集进行实验时,WiLCount 模型的识别精度达到 99.58%,这表明 WiLCount 模型能够有效地适应混合活动状态。

实验结果表明了 WiLCount 模型在不同活动状态下人数识别性能均有良好表现,并且在面对多样化的活动场景时,依然能够保持高水平的准确性,体现了较好的模型鲁棒性。

2) 不同环境对比

为验证不同环境对 WiLCount 模型性能的影响,在小型休息室环境下采集数据,小型休息室的布局与此前的实验环境有较大差异。由于 Wi-Fi 信号易受环境影响,因此在两个环境中采集的数据客观上存在较大差异。原实验环境空间较大,Wi-Fi 感知的范围较大,人员分布分散,相互干扰较小;小型休息室环境空间较小,Wi-Fi 感知的范围较小,人员分布拥挤,相互干扰较大。通过在小型休息室场景下进行实验,能够验证 WiLCount 模型在拥挤环境下的模型识别效果,从而更好地评估模型的鲁棒性与泛化能力。此外,为确保评估的全面性,将保持相同的实验设置和数据采集方法,以便直接对比模型在不同场景中的表现,结果如表 8 所列。

可以看出,在原实验环境中,WiLCOUNT模型的识别精度达到了99.58%。当实验场景切换至小型休息室时,模型的识别精度为97.03%。精度差异产生的原因主要是小型休息室空间较小,Wi-Fi感知范围因此受到限制,人员在该环境中的分布较为密集,导致相互之间的干扰增加。拥挤环境中的信号反射、折射和衍射效应更为复杂,使得信号的多径效应显著增强,从而增加了CSI数据的复杂度。尽管在小型休息室的实验中,模型的识别精度略有下降,但97.03%的准确率仍表明WiLCOUNT在人员密集、干扰较大的环境中具备较好的识别能力。实验结果有效证明了WiLCOUNT模型在应对不同环境时具有较强鲁棒性与良好泛化能力。

表8 不同环境的实验结果

Table 8 Experimental results in different environments

(%)			
环境	准确率	精确率	召回率
数据采集室	99.58	99.61	99.60
小型休息室	97.03	97.29	97.29

3) 不同站立规则对比

为验证不同站立规则下WiLCOUNT模型的性能,在与Count-Fi数据集相同的实验场景中,不限制受测人员站位、行为和行走方向,采集受测者位于任意位置进行无序的随机站立或行走动作数据,将此数据集记录为Random数据集。分别使用Random数据集和Count-Fi数据集进行实验对比,结果如表9所列。

表9 不同站立规则的实验结果

Table 9 Experimental results with different standing rules

(%)			
数据集	准确率	精确率	召回率
Random	97.27	97.22	97.23
Count-Fi	99.58	99.61	99.60

从表9中可以发现,WiLCOUNT模型在使用Count-Fi数据集时的识别准确率为99.58%。当使用更加接近日常生活中真实环境的Random数据集时,WiLCOUNT的识别准确率为97.27%。Random数据集更能反映真实环境的特征,并且相较于Count-Fi数据集,其准确率仅下降2.31%。实验结果表明,WiLCOUNT模型在处理更为复杂的真实环境数据时,依然能够保持较高的识别准确率,体现出其较好的鲁棒性。

综上所述,WiLCOUNT模型不仅在理想环境下表现出色,在更贴近真实环境的情况下也展现出良好的识别性能。这表明WiLCOUNT模型能够在复杂的实际场景中进行精确的人数识别工作。

5.2.4 不同模型对比

目前,Wi-Fi感知人数识别领域并没有可用于模型对比的公开基准模型,为更完整地评估WiLCOUNT模型在人数识别任务中的表现,本文在相同的实验条件下,通过运行深度学习领域的通用复杂模型(Inception-V3^[35], ResNet34^[36], VIT^[37], VanillaNet^[38])、经典轻量化模型(MobileNet-V2^[39], SqueezeNet^[40])和Wi-Fi感知轻量化模型WiDoor^[41]代码进行模型对比实验。选择上述模型的主要依据在于,它们在深度学习领域的广泛应用及其各自的特性。复杂模型具有强大

的特征学习能力和更高的参数容量,通常在各种计算任务中表现优异。通过与这些复杂模型对比,可以验证WiLCOUNT在不依赖复杂网络结构情况下的识别效果,并展示其在保持较小参数量的同时能够超越复杂模型的识别精度。轻量化模型则以减小计算复杂度和减少模型参数为目标,通过将WiLCOUNT与轻量化模型对比,可以评估其在性能与效率之间的平衡能力。此外,由于Wi-Fi感知人数识别领域尚无公开的基准模型,WiDoor作为Wi-Fi感知领域少数公开的轻量化模型,具备较强的代表性。上述模型不仅涵盖了当前深度学习中主流的复杂模型和轻量化模型,还通过与不同结构和机制的模型进行对比,以全面评估WiLCOUNT在人数识别任务中的性能优势。为确保模型对比的公平性和准确性,同时使用振幅和相位两种数据对上述模型进行训练,实验结果如图10所示。

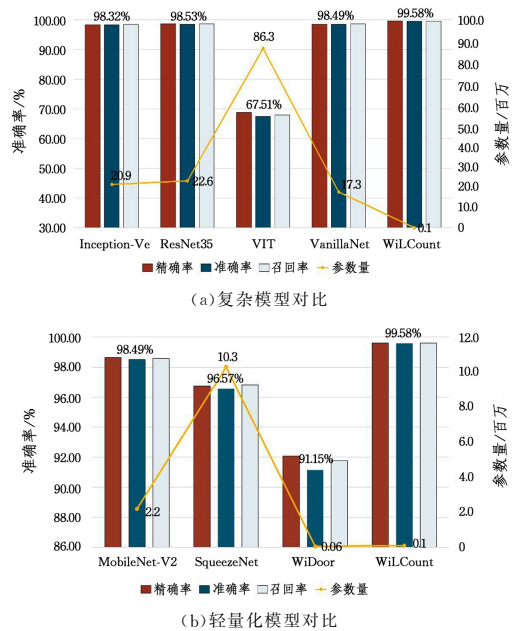


图10 不同模型对比实验

Fig. 10 Comparison experiment of different models

从图10可以看出,Inception-V3, ResNet34, VIT和VanillaNet复杂模型的识别准确率分别为98.32%, 98.53%, 67.51%和98.49%。相比之下,MobileNet-V2, SqueezeNet和WiDoor轻量化模型的识别准确率分别为98.49%, 96.57%和91.15%。由于数据集规模较小,VIT模型的表现不够理想,其余复杂模型以其更多的参数和强大的特征学习能力在识别准确率方面普遍占据优势。而轻量化模型在模型参数规模上仅为复杂模型的1/5,甚至更少。WiLCOUNT模型在人数识别任务中表现最优,其识别准确率高达99.58%。具体而言,WiLCOUNT的准确率与Inception-V3和ResNet34相差约1%,但其参数规模仅为Inception-V3和ResNet34的0.5%。与轻量化模型MobileNet-V2相比,WiLCOUNT在准确率高出1%的同时,参数规模也仅为MobileNet-V2的4%。值得注意的是,尽管WiDoor的参数量相比WiLCOUNT减少了35%,但WiLCOUNT的识别准确率却比WiDoor高9.4%。在兼具识别准确率和轻量化的资源受限场景下,WiLCOUNT具有明显优势。

为深入分析 WiLCount 模型的优越性并验证模型构建的合理性,本文进行了消融实验,结果如表 10 所列。其中,①表示幅相融合模块,②表示单卷积幅相融合模块,③表示 DMConv 模块,④表示单分支 DMConv 模块。

表 10 消融实验结果

Table 10 Results of ablation experiment (%)

模块	准确率	精确率	召回率
①+④	97.51	97.68	97.71
②+③	94.51	95.06	94.94
WiLCount	99.58	99.61	99.60

由表 10 可知,仅使用幅相融合模块+单分支 DMConv 模块或单卷积幅相融合模块+DMConv 模块的实验效果并不理想。首先,对于幅相融合模块+单分支 DMConv 模块的组合,虽然幅相融合模块将 CSI 数据中的振幅和相位的初级特征进行提取和充分融合,显著提升了模型的输入特征多样性,但单分支 DMConv 模块限制了它对不同尺度特征的感知能力,使得它在处理复杂特征时表现不足。尽管幅相融合模块提供了丰富的输入,单分支 DMConv 模块却无法充分利用这些融合后的信息,导致识别效果不佳。其次,对于单卷积幅相融合模块+DMConv 模块的组合,单卷积幅相融合模块仅通过一个卷积操作提取特征。由于幅相信息中蕴含着不同维度的环境变化,而这种简单的操作在融合振幅和相位信息时,只是将两者简单叠加而无法有效融合幅相特征,导致特征提取不完全。虽然 DMConv 模块引入多分支结构来实现不同尺度的特征提取,多尺度特征提取允许模型同时处理不同大小的局部特征,但是单卷积幅相融合模块所提供输入的特征信息不足,即使通过 DMConv 模块提高特征提取的多样性,模型性能也会受限。然而,WiLCount 模型将幅相融合模块与 DMConv 模块结合,其人数识别准确率显著提高至 99.58%。识别效果比仅使用幅相融合模块+单分支 DMConv 模块或单卷积幅相融合模块+DMConv 模块有了较大的提高。幅相融合模块确保了输入特征的多样性,为后续的 DMConv 模块提供了更加全面的特征信息。DMConv 模块则充分发挥其对多尺度特征的提取能力。这种组合有效解决了单一模块在特征提取中的局限性。综上,幅相融合模块为模型提供了更全面的特征输入,而 DMConv 模块则通过多尺度特征提取增强了模型的处理能力,二者结合确保了 WiLCount 模型在复杂场景下的鲁棒性和识别精度。

结束语 传统基于 Wi-Fi 的人数识别模型通常仅利用 CSI 振幅信息而忽略相位信息。相位信息缺失可能导致模型失去对相位变化关键信息的感知,从而限制了在复杂环境中对人数变化的敏感性。针对此问题,本文提出了一种基于 CSI 幅相融合的轻量级模型 WiLCount。WiLCount 通过幅相融合模块完成对振幅和相位信息的充分利用,再使用 DMConv 模块实现对 CSI 特征的全方位、精细化提取。实验结果表明,这两个模块的协同使用可使 WiLCount 在保持高准确率的同时显著减少模型的参数量。

未来研究将聚焦在两个方面:1)研究人体自由活动下的人数识别,从而更好地模拟日常生活中的实际使用情境;2)将尝试提升模型在不同环境下的自适应能力,以便在不同场景

下使用通用模型进行人数识别。

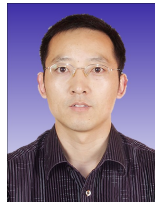
参考文献

- [1] CHOI J W, QUAN X, CHO S H. Bi-Directional Passing People Counting System based on IR-UWB Radar Sensors[J]. IEEE Internet of Things Journal, 2017, 5(2): 512-522.
- [2] LIU P C, NGUANG S K, PARTRIDGE A. Occupancy Inference Using Pyroelectric Infrared Sensors Through Hidden Markov Models[J]. IEEE Sensors Journal, 2016, 16(4): 1062-1068.
- [3] ZHANG Y Y, ZHOU D S, CHEN S Q, et al. Single-Image Crowd Counting via Multi-Column Convolutional Neural Network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 589-597.
- [4] LIN H, MA Z H, JI R, et al. Boosting Crowd Counting via Multifaceted Attention[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2022: 19596-19605.
- [5] LU Y, LYU S H, WANG X D, et al. A survey on WiFi based human behavior analysis technology[J]. Chinese Journal of Computers, 2019, 42(2): 231-251.
- [6] HAN Z J, LU Z M, WEN X M, et al. CentiTrack: Toward Centimeter-Level Passive Gesture Tracking With Commodity WiFi[J]. IEEE Internet of Things Journal, 2023, 10(14): 13012-13027.
- [7] ZHANG C S, JIAO W G. ImgFi: A High Accuracy and Lightweight Human Activity Recognition Framework Using CSI Image[J]. IEEE Sensors Journal, 2023, 23(18): 21966-21977.
- [8] LIU K Z, PEI D S, ZHANG S K, et al. WiCrew: Gait-Based Crew Identification for Cruise Ships Using Commodity WiFi[J]. IEEE Internet of Things Journal, 2023, 10(8): 6960-6972.
- [9] TEWES S, HEINRICHS M, KRONBERGER R, et al. IRS-Enabled Breath Tracking With Colocated Commodity WiFi Transceivers[J]. IEEE Internet of Things Journal, 2023, 10(8): 6870-6886.
- [10] JIN R N, ZHOU J Y, HU J H, et al. Toward Practical Lightweight Passive Human Tracking Using WiFi Sensing[J]. IEEE Internet of Things Journal, 2023, 10(15): 13769-13783.
- [11] BAHL P, PADMANABHAN V N. RADAR: an in-building RF-based user location and tracking system[C]//Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE, 2000: 775-784.
- [12] NAKATSUKA M, IWATANI H, KATTO J. A study on passive crowd density estimation using wireless sensors[C]//In The 4th Intl. Conf. on Mobile Computing and Ubiquitous Networking. 2008.
- [13] YOSHIDA T, TANIGUCHI Y. Estimating the number of people using existing WiFi access point in indoor environment[C]//International Conference on Evolutionary Computing; European Conference of Computer Science. IEEE, 2015.
- [14] DEPATLA S, MURALIDHARAN A, MOSTOFI Y. Occupancy Estimation Using Only WiFi Power Measurements[J]. IEEE Journal on Selected Areas in Communications, 2015, 33(7): 1381-1393.
- [15] LI H C, CHAN E, GUO X N, et al. Wi-Counter: Smartphone-

- Based People Counter Using Crowdsourced Wi-Fi Signal Data [J]. *IEEE Transactions on Human-Machine Systems*, 2015, 45(4):442-452.
- [16] KURA S, YAMAGUCHI H, SHIRAIISHI Y. Low-cost Pedestrian Counter Using Wi-Fi APs for Smart Building Applications [C]// 2018 IEEE 42nd Annual Computer Software and Applications Conference. IEEE, 2018:640-645.
- [17] HALPERIRR D, HU W, SHETH A, et al. Tool release: gathering 802.11n traces with channel state information [J]. *ACM SIGCOMM Computer Communication Review*, 2011, 41(1):53-53.
- [18] YANG Z, ZHOU Z M, LIU Y H. From RSSI to CSI: Indoor Localization via Channel Response [J]. *ACM Computing Surveys*, 2013, 46(2):1-32.
- [19] XI W, ZHAO J Z, LI X Y, et al. Electronic frog eye: Counting crowd using WiFi [C]// IEEE INFOCOM 2014 - IEEE Conference on Computer Communications. IEEE, 2014:361-369.
- [20] DING Y S, GUO B, XIN T, et al. WiCount: A Crowd Counting Method Based on WiFi Channel State Information [J]. *Computer Science*, 2019, 46(11):297-303.
- [21] WANG S N, XUN Y J, ZHAO J, et al. A Novel Personnel Counting Method Based on WiFi Perception [C]// 2022 IEEE 23rd International Conference on High Performance Switching and Routing. IEEE, 2022:206-211.
- [22] CHOI H, FUJIMOTO M, MATSUI T, et al. Wi-CaL: WiFi Sensing and Machine Learning based Device-Free Crowd Counting and Localization [J]. *IEEE Access*, 2022, 10:24395-24410.
- [23] GUO Z X, FU X, SHENG B Y, et al. TWCC: A Robust Through-the-Wall Crowd Counting System Using Ambient WiFi Signals [J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(4):4198-4211.
- [24] HOU H W, BI S Z, ZHENG L L, et al. DASECount: Domain-Agnostic Sample-Efficient Wireless Indoor Crowd Counting via Few-Shot Learning [J]. *IEEE Internet of Things Journal*, 2022, 10(8):7038-7050.
- [25] KHAN D, HO I. CrossCount: Efficient Device-Free Crowd Counting by Leveraging Transfer Learning [J]. *IEEE Internet of Things Journal*, 2023, 10(5):4049-4058.
- [26] YOUSSEF M, MAH M, AGRAWALA A. Challenges: Device-free passive localization for wireless environments [C]// ACM/IEEE International Conference on Mobile Computing and Networking. ACM, 2007:9-14.
- [27] BAO N, DU J J, WU C Y, et al. Wi-Breath: A WiFi-Based Contactless and Real-Time Respiration Monitoring Scheme for Remote Healthcare [J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 27(5):2276-2285.
- [28] SEN S, RADUNOVIC B, CHOUDHURY R R, et al. You are facing the Mona Lisa: spot localization using PHY layer information [C]// ACM SIGMOBILE International Conference on Mobile Systems, Applications, and Services. ACM, 2012:183-196.
- [29] QIAN K, WU C S, YANG Z, et al. PADS: Passive detection of moving targets with dynamic speed using PHY layer information [C]// 2014 20th IEEE International Conference on Parallel and Distributed Systems. IEEE, 2014:1-8.
- [30] LIU Z J, WANG L, LIU W Y, et al. Human Movement Detection and Gait Periodicity Analysis Using Channel State Information [C]// International Conference on Mobile Ad-hoc & Sensor Networks. IEEE, 2016:167-174.
- [31] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [J]. arXiv:1704.04861, 2017.
- [32] XIE Y X, LI Z J, LI M. Precise Power Delay Profiling with Commodity Wi-Fi [J]. *IEEE Transactions on Mobile Computing*, 2015, 18:1342-1355.
- [33] GRINGOLI F, SCHULZ M, LINK J, et al. Free Your CSI: A Channel State Information Extraction Platform For Modern Wi-Fi Chipsets [C]// 13th ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization. ACM, 2019:21-28.
- [34] HUSSEIN K, CLAUDIO P, FEDERICO P, et al. Impact of Wi-Fi traffic on the IEEE 802.15.4 channels occupation in indoor environments [C]// 2009 International Conference on Electromagnetics in Advanced Applications. IEEE, 2009:1042-1045.
- [35] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the Inception Architecture for Computer Vision [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016:2818-2826.
- [36] HE K M, ZHANG X, REN S Q, et al. Deep Residual Learning for Image Recognition [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015:770-778.
- [37] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale [J]. arXiv:2010.11929, 2020.
- [38] CHEN H T, WANG Y H, GUO J Y, et al. VanillaNet: the Power of Minimalism in Deep Learning [J]. arXiv:2305.12972, 2023.
- [39] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018:4510-4520.
- [40] IANDOLA F N, MOSKEWICZ M W, ASHRAF K, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1 MB model size [J]. arXiv:1602.07360, 2016.
- [41] CAO C Y, YANG X D, DUAN P S. WiDoor: Close-range Contactless Human Identification Approach [J]. *Computer Science*, 2023, 50(4):388-396.



DUAN Pengsong, born in 1983, Ph.D, associate professor, is a member of CCF (No. 43410M). His main research interests include wireless sensing, IoT and machine learning.



CAO Yangjie, born in 1976, Ph.D, professor, is a member of CCF (No. 17620S). His main research interests include machine learning, computer vision and high-performance computing.