

## 基于CSI实例标准化的域泛化人体动作识别模型

王杨<sup>1</sup>, 许佳炜<sup>1</sup>, 王傲<sup>1</sup>, 夏慧娟<sup>1</sup>, 赵传信<sup>1</sup>, 季一木<sup>2</sup>

(1.安徽师范大学计算机与信息学院, 安徽 芜湖 241002; 2.南京邮电大学计算机学院, 江苏 南京 210003)

**摘要:** 为了实现完全不依赖目标域数据的Wi-Fi跨域人体动作感知, 提出了一种基于CSI实例标准化的域泛化人体动作识别模型INDG-Fi。INDG-Fi使用实例标准化去除CSI特征表示的领域信息, 接着构建共享特征提取的动作分类器和域分类器, 并通过动作偏向学习和对抗性的域学习, 将编码层提取的特征偏向人体动作引起的信号特征, 同时远离领域信号影响。为了让模型关注受人体动作影响更显著的子载波信号, 在编码层中加入子载波注意力模块。实现结果表明, 所提INDG-Fi在不可见的用户和位置的感知性能分别为97.99%和92.73%, 能够实现鲁棒的跨域感知。

**关键词:** 信道状态信息; 无线感知; 人体动作识别; 域泛化

**中图分类号:** TP391

**文献标志码:** A

**DOI:** 10.11959/j.issn.1000-436x.2024110

## Domain-generalization human activity recognition model based on CSI instance normalization

WANG Yang<sup>1</sup>, XU Jiawei<sup>1</sup>, WANG Ao<sup>1</sup>, XIA Huijuan<sup>1</sup>, ZHAO Chuanxin<sup>1</sup>, JI Yimu<sup>2</sup>

1. School of Computer and Information, Anhui Normal University, Wuhu 241002, China

2. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

**Abstract:** To achieve Wi-Fi cross-domain human activity perception that was not dependent on target domain data, a domain-generalization human activity recognition model based on CSI instance normalization called INDG-Fi was proposed. The instance normalization standardization was utilized to remove domain information from the representation of CSI features by INDG-Fi. Then action classifiers and domain classifiers were constructed for shared feature extraction. By employing activity bias learning and adversarial domain learning, the model biased the features extracted from the encoding layer towards signal variations caused by human actions while moving away from domain signals. To enhance the model's focus on subcarrier signals that were more significantly influenced by human actions, a subcarrier attention module was incorporated into the encoding layer. The implemented results demonstrate that the proposed INDG-Fi achieves perceptual accuracies of 97.99% and 92.73% for unseen users and locations, respectively, thus enabling robust cross-domain perception.

**Keywords:** channel state information, wireless sensing, human activity recognition, domain-generalization

收稿日期: 2024-03-27; 修回日期: 2024-05-28

通信作者: 王杨, wycap@126.com

基金项目: 国家自然科学基金资助项目(No.61871412); 江苏省重点研发计划基金资助项目(No.BE2023004-2); 安徽省自然科学基金重点项目(No.KJ2019A0938, No.KJ2021A1314, No.KJ2019A0979)

**Foundation Items:** The National Natural Science Foundation of China (No.61871412), The Key Research and Development Program of Jiangsu Province (No.BE2023004-2), The Natural Science Foundation Key Project of Anhui Province (No.KJ2019A0938, No.KJ2021A1314, No.KJ2019A0979)

### 0 引言

基于 Wi-Fi 的无线感知技术已经在活动识别<sup>[1-2]</sup>、手势识别<sup>[3]</sup>、身份认证<sup>[4]</sup>、室内定位<sup>[5]</sup>和运动跟踪<sup>[6]</sup>领域被广泛应用。然而, Wi-Fi 信号在传播过程中容易受到反射、衍射或散射的影响,对环境变化异常敏感。此外,由于个体体征和行为习惯的差异,不同用户完成同一动作时所引起的信道状态信息(CSI, channel state information)幅值变化也存在较大不同<sup>[6]</sup>,如图1所示。研究者通常使用“域”来描述与人体动作无关的因素(如环境和用户),在实际应用中由于源域和目标域的分布差异,将在源域数据集上训练的 Wi-Fi 感知模型应用于新的目标域时,模型准确率明显降低。

为了解决 Wi-Fi 跨域感知的问题,学术界提出了多种解决方案<sup>[7-10]</sup>,主要包括基于域无关特征、基于域对抗训练、基于迁移学习和基于小样本学习等方法。基于域无关特征的方法通常需要手动设计特征,且依赖于多个特征提取器。其任务是从输入数据中提取出具有代表性的特征表示,当接收器数

量不足时, Wi-Fi 跨域感知性能明显下降。基于域对抗训练的方法是通过特征提取器和域分类器之间进行相互对抗学习。其可以促使特征提取器学习到具有泛化能力的特征表示,从而在新的目标域上有更好的性能表现。但目前基于域对抗训练的方法依赖大量来自源域和目标域的样本进行训练。基于迁移学习的方法旨在利用预训练模型,将源域的知识迁移到新的目标域上。然而,在实际场景下,当源域和目标域之间存在较大差异时,可能会出现负迁移现象。基于小样本学习的方法利用少量目标域数据就可以将源域映射到目标域,并且不需要重新训练模型。但小样本学习器并不能学习到域无关的特征表示,仅能在基本集的多个任务下找到最优的起点,以便快速地适应测试任务。

本文旨在增强基于 Wi-Fi 跨域感知的能力,以更好地满足实际应用场景的需求,即使在面对完全不可见的目标域时,也可实现鲁棒的跨域感知。假设用  $X$  表示输入空间,  $Y$  表示对应的标签输出,源域  $S = \{(x_i, y_i)\}_{i=1}^n \sim P_{XY}$ 。本文模型旨在从多个源域

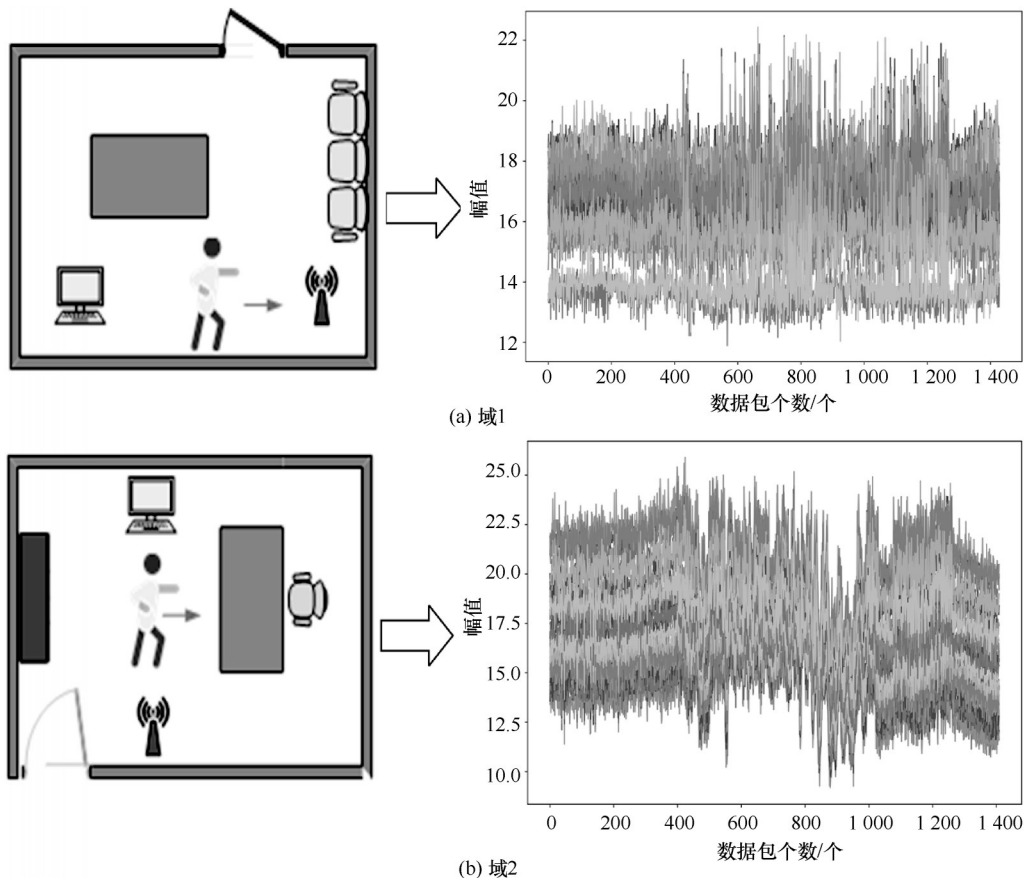


图1 同一动作在不同域下的CSI幅值变化

$S = \{S_i\}_{i=1}^m$  ( $m$  表示源域个数) 训练得到预测函数  $g: X \rightarrow Y$ , 该函数可以学习到泛化的特征表示, 在不可见的目标域  $T$  上预测损失也较低。

为了实现这一目标, 本文设计出带有子载波注意力 (SA, subcarrier attention) 模块的编码层将 CSI 矩阵映射到特征表示  $z$ , 并使用统计量均值  $u(z)$  和方差  $\sigma(z)$  作为子载波的域样式。在本文模型训练过程中, 加入动作偏向学习和对抗性的域学习。在动作偏向学习任务中, 通过对均值和方差进行随机插值, 得到模糊子载波的域样式, 接着应用视觉领域风格迁移的自适应实例归一化 (AdaIN, adaptive instance normalization) 算法实现子载波的风格迁移, 以此让编码层提取 CSI 特征时远离领域信号, 让模型学习到人体动作引起的域无关子载波信号变动, 如图 2 所示。对抗性域动作偏向学习加入对抗性训练的的目的是让共享编码层进一步远离领域信号的干扰。

综上所述, 本文模型在对 CSI 实例标准化基础上, 分别使用域随机化和风格迁移, 以此实现域泛化的 Wi-Fi 跨域感知模型, 主要研究工作如下。

- 1) 从更具有挑战性的域泛化学习方式出发, 本文所提 INDG-Fi 感知模型即使在面对目标域不可见情景时, 也能实现鲁棒的跨域感知。
- 2) 设计动作偏向学习和对抗性的域学习, 同时在编码层中加入子载波注意力模块, 使 INDG-Fi 模型能更好地学习到域无关的动作信号。
- 3) 在真实场景中进行了广泛的跨域感知实验, 以此验证本文模型对比基线模型的性能提升。

## 1 相关工作

无线信号经历多径传播, 其特征由环境和传感活动共同决定。在不同环境中执行相同活动将导致不同的 CSI 变化。在实际场景下, 不同环境或不同用户中应用 Wi-Fi 跨域感知是不可避免的。现有的跨域感知解决方案主要有基于域无关特征的方法和基于域自适应学习的方法<sup>[11]</sup>。

先前研究通常使用特定的技术在训练模型之前推导出与位置无关的特征。Widar3.0 推导出域无关的体坐标速度剖面 (BVP, body-coordinate velocity profile) 特征可实现 “Zero-Effort” 的跨域手势识别, 但 BVP 需要 6 个发送-接收链路进行计算且计算复杂度高<sup>[9]</sup>。WiHF 开发出双任务深度神经网络 (DNN, deep neural network) 框架, 利用运动模式作为位置无关的特征, 协同识别手势和用户, 设备需求与 Widar3.0 相同。WiGesture 设计了一种新的位置无关传感策略, 并以手势识别作为应用实例来验证该策略的有效性, 将视角从传统的收发器转向面向手势的角度, 提取出位置无关的特征动作导航原语 (MNP, motion navigation primitive)<sup>[12]</sup>。Niu 等<sup>[13]</sup>通过分析 Wi-Fi 传感中的多普勒频移现象, 构建了 Wi-Fi 频率模型, 用于量化信号频率与人类活动目标的位置、运动方向和速度之间的关系, 将运动片段和相对运动方向变化识别为 2 个位置无关的特征。Gao 等<sup>[14]</sup>将每个手势的信号时间序列根据其质量进行细分, 使用特定的信号处理技术分别进行处理, 通过建立一个将手势信号与环境噪声联系起来的数学模型来表征信号的感知质量, 由此推导出动态相位指数误差, 用于定量描述每个手势信号的

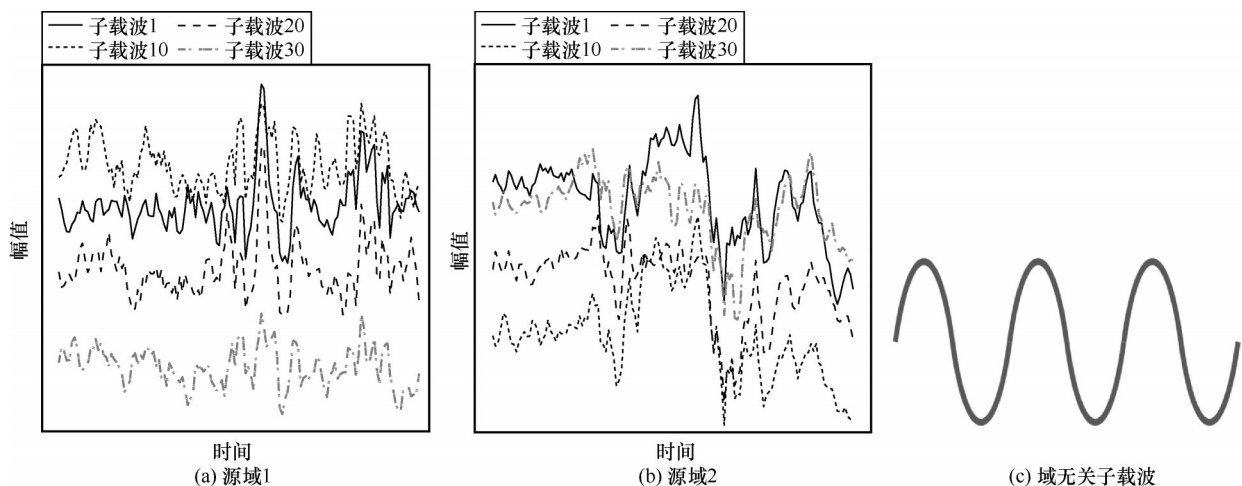


图 2 不同域下人体动作对子载波的影响

感知质量。

基于域无关特征的方法不仅要求设计特征时具备专业的领域知识,还依赖多个接收器来收集数据,并且对传感区域有一定的限制。因此,基于域无关特征的方法增加了 Wi-Fi 跨域感知成本,与其低成本的特性相矛盾。基于域自适应学习的方法不依赖人工设计的域无关特征,利用深度神经网络提取复杂特征。Sheng 等<sup>[15]</sup>提出了一个深度学习框架,为解决训练模型随着环境变化而完全失效的问题,将深度学习模型作为预训练模型,并在新场景中对其进行微调,该方法能够以较低的计算量和较好的准确率实现跨场景动作识别。Bu 等<sup>[16]</sup>同样使用迁移学习的方法实现跨域手势识别。Jiang 等<sup>[17]</sup>提出了环境独立 (EI, environment independent) 感知框架,通过域对抗训练方式去除活动数据中包含的环境信息,提取出不同环境下的域无关特征。Kang 等<sup>[18]</sup>将对抗学习方案与特征解纠缠模块相结合,为了促进源域到目标域的正向迁移并消除与手势信号无关的信息,设计源域和目标域两阶段对齐的策略以及基于源域和目标域的判别信息的注意力方案。Zou 等<sup>[19]</sup>提出了一种无监督联合对抗域自适应 (JADA) 模型,该模型设计了 4 个阶段的学习策略: 1) 使用标记的源域数据训练源编码器和动作分类器; 2) 训练目标编码器并通过无监督的域对抗训练方式使域分类器无法区分数据的域标签; 3) 再次使用标记的源域数据重新训练源编码器和共享动作分类器; 4) 利用阶段 2) 的目标编码器和阶段 3) 的共享动作分类器在测试数据集上实现跨域感知。

上述方法必须获得足量的目标域数据来训练和

评估模型,以实现模型从源域到目标域的自适应,在实际场景应用时难免耗费人力资源。因此,只需要少数目标域数据即可将源域映射到目标域的小样本学习方式引起了学者关注。Shi 等<sup>[20]</sup>提出了有效地利用之前看到的环境 (PSE, previously seen environment) 数据集,该训练策略可以有效地实现人类活动识别。每个活动仅使用来自测试环境的一个样本和一个 PSE 数据集,训练时间更短且识别准确率更高。Zhou 等<sup>[21]</sup>为充分利用所有数据,提出了一种面向目标的半监督 (TOSS, target-oriented semi-supervised) 域自适应方法,可以有效地利用标记和未标记的目标样本,并设计了动态伪标签策略和基于不确定性的选择方法以便于从源环境和目标环境中学习知识。上述工作虽然缓解了目标域中数据收集的负担,但每个目标域训练模型的成本是昂贵的,并且有些场景下收集所有潜在域数据是不可能的。

## 2 系统模型

INDG-Fi 模型系统框架如图 3 所示。首先从源域数据集中采样 CSI 样本  $x_{ori}$ , 维度为  $(N_t, N_r, K, T)$ , 其中,  $N_t$  表示发射端天线数量,  $N_r$  表示接收端天线数量,  $K$  表示每对天线之间的子载波个数,  $T$  表示采样的序列长度, 经过数据预处理模块得到  $x$ , 维度为  $(N_t, N_r, K, T')$ , 其中,  $T'$  表示对齐后采样的序列长度, 接着由编码层将  $x$  映射为特征  $z \in \mathbb{R}^d$ , 计算出  $z$  的均值  $u(z)$  和方差  $\sigma(z)$ 。然后在动作偏向学习任务中, 特征  $z$  先使用域模糊化得到  $\hat{u}(z, z')$  和  $\hat{\sigma}(z, z')$ , 再利用 AdaIN 进行去风格化处理, 接着输入动作分类器中, 构建手势分类损失  $L_g$ 。在对抗性的域学习任务中, 直接使用 AdaIN 对  $z$  进行去风格化处理,

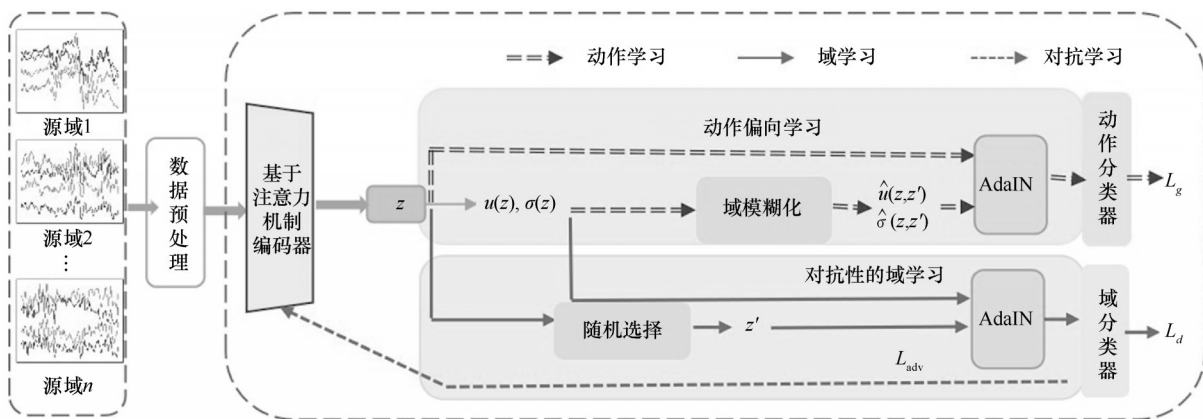


图3 INDG-Fi 模型系统框架

构建域分类损失  $L_d$ ，同时加入对抗损失  $L_{adv}$  使编码器进一步远离领域信号的影响。

### 2.1 编码层网络

编码器网络通过一维卷积学习不同子载波空间上的特征关系并不断对 CSI 数据进行时间尺度上的压缩，最后展平为特征  $z$ 。编码层网络主要组成有编码块和聚合块，其结构如图 4 所示。其中，编码块包含卷积 (Conv) 层、批归一化 (BatchNorm) 层、非线性激活 (ReLU) 层、最大池化 (MaxPool) 层和子载波注意力 (SubCarrier Attention) 层。

由于模型输入为全部接收端天线的子载波，而人体动作会对不同子载波产生不同的影响。因此，区别于基线模型采用残差块，本文所提感知模型设计子载波注意力层来表征不同子载波的重要程度。通过在子载波上增加注意力模块帮助模型更好地学习到最相关的子载波，颜色越深表示对人体动作子载波影响越大，模型对其关注度也更高。为方便后续算法介绍，这里给出编码层的处理过程：由数据预处理得到  $x$ ，经过编码器网络得到特征序列  $z = f_{encoder}(x)$ 。

### 2.2 动作偏向学习

动作偏向学习的任务旨在让本文所提模型学习到不受域影响的人体动作信号。使用  $C_{\phi_g}$  表示动作分类器，目标是学习到参数  $\phi_g$ ，使得动作分类器在该参数下分类时不受域信号的影响。为了实现这一目标，本文对均值  $u(z)$  和方差  $\sigma(z)$  进行域模糊化处理。具体而言，随机从一个小批次的训练样本中

选择  $x'(x \neq x')$ ，由编码器得到  $z'$ 。为了让模型更能学习到鲁棒的特征表示，首先  $z'$  的选择尽量不固定，然后通过线性插值的方式得到域模糊化的  $\hat{u}$  和  $\hat{\sigma}$ ，分别如式(1)和式(2)所示。

$$\hat{\mu}(z, z') = \alpha\mu(z) + (1 - \alpha)\mu(z') \quad (1)$$

$$\hat{\sigma}(z, z') = \alpha\sigma(z) + (1 - \alpha)\sigma(z') \quad (2)$$

其中， $\alpha \sim \text{Uniform}(0,1)$  为线性插值的权重。

其次，域模糊化后的  $\hat{u}$  和  $\hat{\sigma}$  通过 AdaIN 替换  $z$  的原始子载波样式，此过程为域随机化 (DR, domain randomization)。

$$\text{DR}(z, z') = \hat{\sigma}(z, z') \left( \frac{z - \mu(z)}{\sigma(z)} \right) + \hat{\mu}(z, z') \quad (3)$$

接着，将  $\text{DR}(z, z')$  混合后的特征输入  $C_{\phi_g}$  中，动作偏向学习的整体损失如式(4)所示。

$$L_g = -E_{(x,y,d) \in S} \sum_{k=1}^{K_g} y_k \log C_{\phi_g}(\text{DR}(z, z')) \quad (4)$$

其中， $K_g$  表示全部的动作类别数， $y_k$  表示输入样本  $x$  所对应的真实标签。

### 2.3 对抗性的域学习

为进一步提高编码层网络学习到与样本所处领域相关特征的能力，增加域学习和对抗学习使特征提取器削弱域无关信号的影响。设  $C_{\phi_d}$  表示域分类器，目标是学习到参数  $\phi_d$ ，使得域分类器在该参数下分类时能有效鉴别出样本所处的域信息。为了让编码层网络更能关注到均值和方差为特征域风格表示，这里域学习任务为辨别域风格迁移后的特征所对应的域标签。首先，随机选择  $x'$ ，并由编码层得

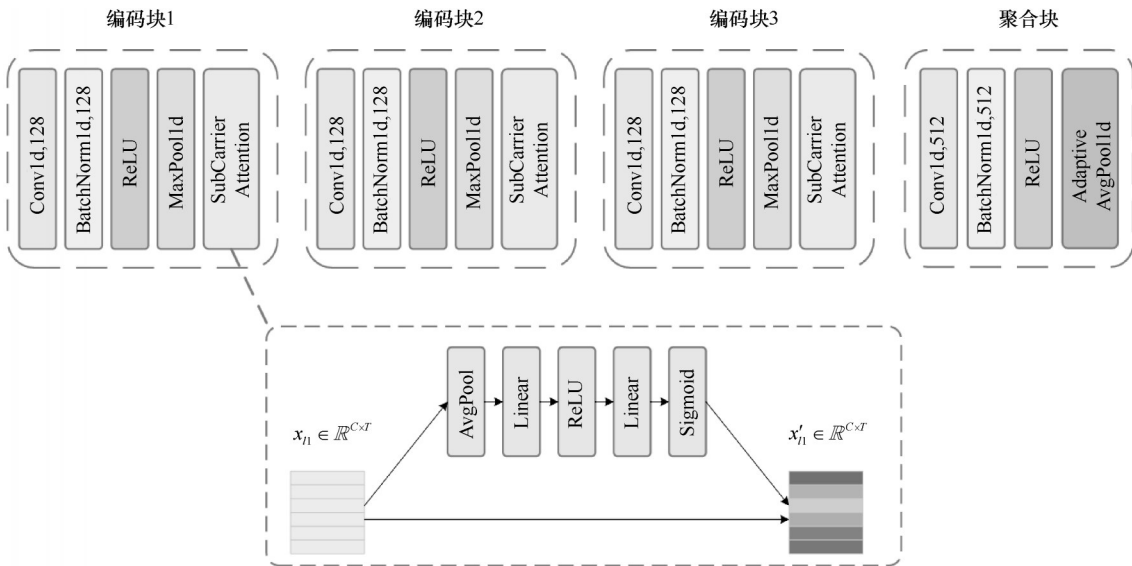


图 4 编码层网络结构

到 $z'$ , 对特征表示 $z'$ 使用AdaIN实现域风格迁移。

$$\text{AdaIN}(z, z') = \sigma(z) \left( \frac{z' - \mu(z')}{\sigma(z')} \right) + \mu(z) \quad (5)$$

接着将风格迁移后的 $z$ 输入 $C_{\phi_d}$ 中, 构建域分类损失, 如式(6)所示。

$$L_d = -E_{(x,y,d) \in \sum_{m_d=1}^{M_d} d_{m_d}} \log C_{\phi_d}(\text{AdaIN}(z, z')) \quad (6)$$

其中,  $M_d$ 表示源域数量, 这里 $z'$ 已经迁移为 $z$ 的域风格, 因此 $d_{m_d}$ 表示输入样本 $x'$ 所对应的域标签。

本文模型采用对抗学习以期学习到域无关的特征表征, 而上述加入域的分类与最终学习目标相矛盾。为解决该问题, 加入对抗学习可以让编码器学习到的特征远离领域信号的影响。

$$L_{\text{adv}} = -\lambda E_{(x,y,d) \in \sum_{k=1}^{K_d} \frac{1}{K_d}} \log C_{\phi_d}(\text{AdaIN}(z, z')) \quad (7)$$

其中,  $\lambda$ 为损失的权重因子。

#### 2.4 算法整体流程

输入多个源域数据集 $S = \{S_i\}_{i=1}^m$ , INDG-Fi模型经过动作偏向学习和对抗性的域学习任务学习能够从中学到域无关的表征, 其训练过程可由算法1表示。

##### 算法1 INDG-Fi训练过程

输入 源域数据集, 编码层网络 $G_{\theta}$ , 动作分类器 $C_{\phi_g}$ , 域分类器 $C_{\phi_d}$

输出 参数更新后的编码层网络 $G_{\theta}$

- 1) for sampled  $\{x_i\}_{i=1}^B$  in  $S$  do
- 2)  $J \leftarrow 0$
- 3) for all  $i \in \{1, \dots, B\}$  do
- 4)  $z_i \leftarrow G_{\theta}(x_i)$
- 5)  $z'_i \leftarrow G_{\theta}(\text{RandomSample}(\{x_i\}_{i=1}^B))$
- 6) 计算  $u(z_i)$ ,  $\sigma(z_i)$ ,  $\mu(z'_i)$ ,  $\sigma(z'_i)$
- 7)  $\hat{\mu}(z_i, z'_i) \leftarrow \alpha \mu(z_i) + (1 - \alpha) \mu(z'_i)$
- 8)  $\hat{\sigma}(z, z') \leftarrow \alpha \sigma(z) + (1 - \alpha) \sigma(z')$
- 9)  $\text{DR}(z_i, z'_i) \leftarrow \hat{\sigma}(z_i, z'_i) \left( \frac{z - \mu(z_i)}{\sigma(z_i)} \right) + \hat{\mu}(z_i, z'_i)$
- 10) 通过式(4), 计算 $L_g$
- 11)  $\text{AdaIN}(z_i, z'_i) = \sigma(z_i) \left( \frac{z'_i - \mu(z'_i)}{\sigma(z'_i)} \right) + \mu(z_i)$
- 12) 通过式(6), 计算 $L_d$
- 13) 通过式(7), 计算 $L_{\text{adv}}$
- 14)  $J \leftarrow L_g + L_d + \lambda L_{\text{adv}}$

15) minimize  $J$  to update  $G_{\theta}$

16) end for

17) end for

### 3 实验设计

#### 3.1 数据采集

任何环境的变化都会对Wi-Fi信号的多径传播轮廓造成影响, 本文的跨域感知模型是在这些变化的源域中学习域无关的动作特征表示。为了验证INDG-Fi模型在真实场景下的跨域感知性能, 本文基于Intel 5300网卡平台采集了多个场景的数据, 并使用CSI-Tool工具套件提取和分析CSI<sup>[22]</sup>。实验使用一台商用小米路由器作为发射端, 2台装有5300网卡的台式电脑作为接收端(后续INDG-Fi模型跨域实验以使用接收端1数据为主, 但也会给出与使用2个接收端的实验结果对比), 终端系统为Ubuntu12.04, 网卡工作频段为2.4 GHz, 采样频率为1 kHz, 每种活动采集的持续时间为2 s。数据采集的场景如图5所示, 其中, 场景1为会议室, 场景2为办公室, 场景3为实验室, TX表示发射端, RX<sub>1</sub>和RX<sub>2</sub>表示2个接收端。

设备具体摆放和采集位置如图6所示, 其中, TX表示发射端, RX<sub>1</sub>和RX<sub>2</sub>表示2个接收端, P<sub>1</sub>~P<sub>4</sub>表示4个不同采集点。接收端1距离发射端2 m, 接收端2在发射端和接收端1中间位置正上方1 m处。用户1~用户4在图中P<sub>1</sub>~P<sub>4</sub>这4个点采集, 并且采集时朝向发射端的3个方向, 如图6中P<sub>2</sub>位置所示的3个箭头。最后在上述环境下, 4名用户采集了6种人体动作, 包括向外推手臂、摆动手臂、举起手臂、鼓掌、画圈和握拳。去除掉丢包过多的数据样本, 本文最终在场景1、场景2和场景3中采集的样本数分别为5 281、5 288和5 300个, 每名用户在每个域上采集样本的平均数量为18个。

#### 3.2 数据预处理

收集到的CSI原始测量结果往往受到噪声的干扰, 不能直接用于模型训练, 因此本文对原始CSI数据分别进行幅值降噪和相位校准<sup>[23]</sup>。1) 本文首先使用Hampel滤波对原始CSI的幅值进行了滤波和平滑处理, Hampel滤波能有效地处理异常点和偶发的大噪声, 同时更好地保留原始信号的特征信息。2) 对于相位信息使用线性相位去噪, 以保留原始CSI信号频率特性, 减小失真。

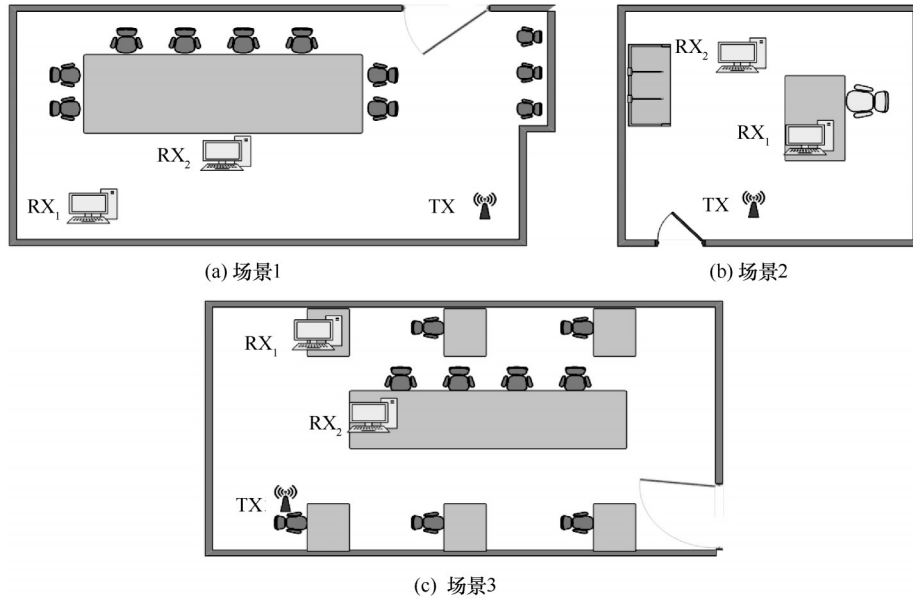


图5 数据采集的场景

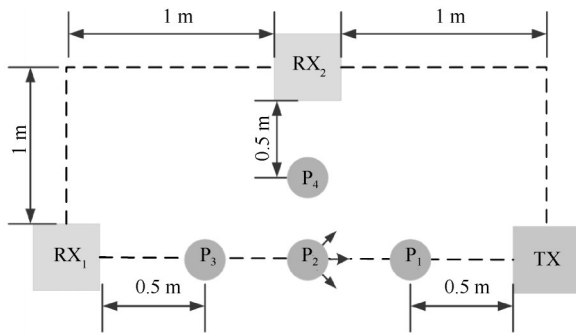


图6 设备具体摆放和采集位置

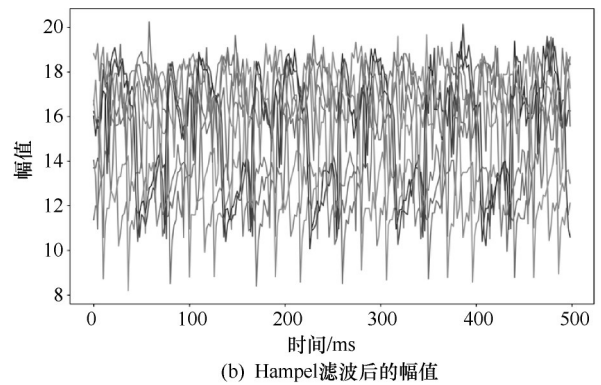
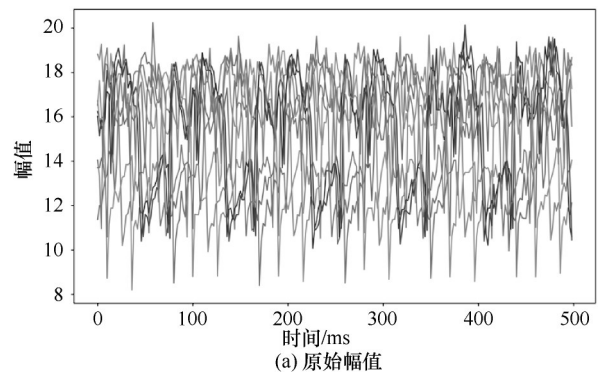


图7 Hampel滤波对子载波幅值的影响

Hampel 滤波是基于滑动窗口的机制对 CSI 序列进行处理，先计算窗口内数据的中位数和标准差，再根据阈值比例数乘标准差得到正常点范围。对于 CSI 序列  $x = [x_1, x_2, \dots, x_N]$ ，长度为  $N$ ，设滑动窗口大小为  $W$ ，窗口中心位置为  $i$ ， $\tilde{x}_i$  表示窗口中位数， $S_i$  为窗口标准差，正常点范围如式(8)所示。

$$[\tilde{x}_i - kS_i, \tilde{x}_i + kS_i] \quad (8)$$

其中， $k$  表示阈值比例。

Hampel 滤波对子载波幅值的影响如图 7 所示，本文将窗口大小设为 5，标准差设为 3，最大程度地保留原始 CSI。从图 7 可以看出，Hampel 滤波没有对原始信号的其他部分造成太大的扭曲，但对一些偶发噪声和离群点进行了有效去噪。因为本文模型需要学习子载波的风格特征，使用 Hampel 滤波不会过多破坏子载波的风格特征。

CSI 相位纠正可以使不同天线之间的信号保持同步，以确保测量结果的准确性。测量到的第  $i$  个子载波的相位为

$$\tilde{\theta}_i = \theta_i + \frac{2\pi K_i \Delta t}{N} + \beta + Z_f \quad (9)$$

其中， $\theta_i$  表示真实相位值， $\Delta t$  表示时延， $\beta$  表示可

能出现的随机相位偏差,  $Z_f$ 表示测量过程中的随机噪声,  $K_i$ 表示子载波索引,  $N = 64$ 表示快速傅里叶变换(FFT, fast Fourier transform)点数。

对于接收器而言,在传输过程中,式(9)中间的两项噪声保持不变,线性相位去噪引入仅与真实相位相关,去噪后的相位为

$$\hat{\theta}_i = \tilde{\theta}_i - \frac{\tilde{\theta}_n - \tilde{\theta}_1}{K_N - K_1} K_i - \frac{1}{N} \sum_{i=1}^N \tilde{\theta}_i \quad (10)$$

由图8可知,线性相位去噪后的每条子载波的变化更能反映最初真实的相位变化信息。

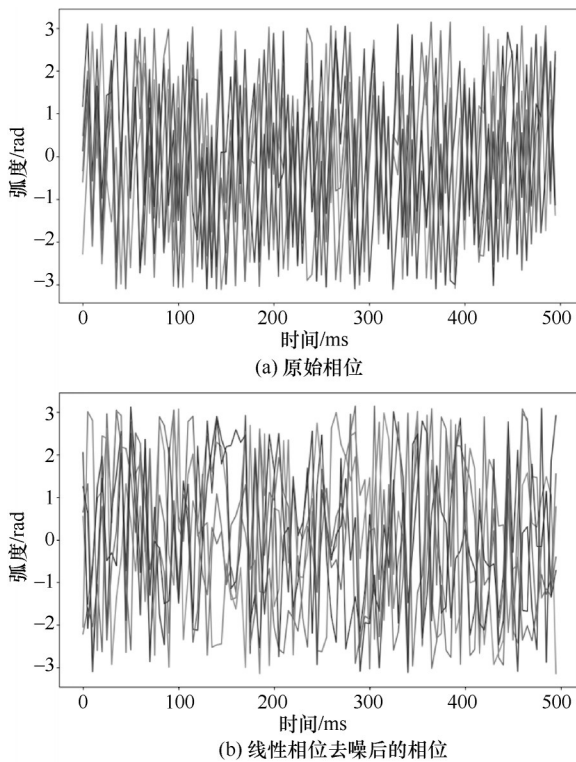


图8 线性相位去噪对子载波的影响

### 3.3 实验设置

本文训练模型的软硬件环境为: Ubuntu 20.04 LTS 操作系统, NVIDIA RTX 3080Ti GPU, Torch1.8.1, cudnn11.1。手势分类、域分类和对抗训练均使用随机梯度下降(SGD, stochastic gradient descent)优化器,学习率为 $5 \times 10^{-3}$ ,动量为0.9,训练迭代次数为200,批次大小为32,对抗损失 $\lambda = 0.2$ 。

本文后续实验均为跨域实验,在单次跨域实验下,训练样本为全部源域样本,测试样本为全部目标域样本。以跨用户(C-U, cross-user)实验为例:若目标域选择用户4的数据,则训练样本为该场景下用户1、用户2和用户3的全部数据。

## 4 实验结果及分析

### 4.1 跨域感知实验

本文数据集收集了3个场景和4名用户在4个不同位置上的人体动作数据。本节给出INDG-Fi模型在这3个场景下跨位置(C-L, cross-location)和跨用户的表现。实验结果如表1所示。在平均准确率上,INDG-Fi模型在所有场景中的表现都相对较好,平均准确率在0.716 1到0.815 7之间,这表明该模型在目标域完全不可见的情况下,仍然具有良好的域泛化能力,能够在不同的场景下保持相对稳定的感知性能。在最高准确率上,在所有场景和跨域类型中,模型的性能达到了相对较高的水平,最高准确率在0.848 4到0.979 9之间,意味着INDG-Fi模型在面对某些未知用户和位置时已经实现精准的自适应。考虑到本文实验是6分类的实验,最低准确率在0.589 0到0.690 5之间,说明虽然INDG-Fi模型某些场景下存在一定性能下降,但是模型仍然可以在多个源域中学习到的不变的特征表示,即模型是具备一定的泛化能力。

表1 3个场景下INDG-Fi模型跨域识别性能

场景	类型	平均准确率	最高准确率	最低准确率
场景1	跨用户	0.716 1	0.848 4	0.555 6
	跨位置	0.815 7	0.932 9	0.690 5
场景2	跨用户	0.767 8	0.927 3	0.577 3
	跨位置	0.735 7	0.893 7	0.589 0
场景3	跨用户	0.812 2	0.910 9	0.640 4
	跨位置	0.809 9	0.979 9	0.617 6

图9给出了3个场景下不同用户的跨位置识别性能,图中的横坐标代表目标域。以横坐标位置1为例,其代表了源域训练数据为位置2、位置3和位置4,预测的目标域为位置1,且依次对每名用户进行跨域感知的实验。因为不同用户在体型和行为习惯上存在差异,所以INDG-Fi模型在不同用户的实验结果中也是有差异,这里以场景1的实验结果为例,用户2在4个位置上整体表现最优,其跨域准确率分别为0.787 4、0.869 6、0.932 9和0.904 8,而用户4表现最差,其跨域准确率分别为0.716 7、0.702 7、0.795 3和0.787 0。

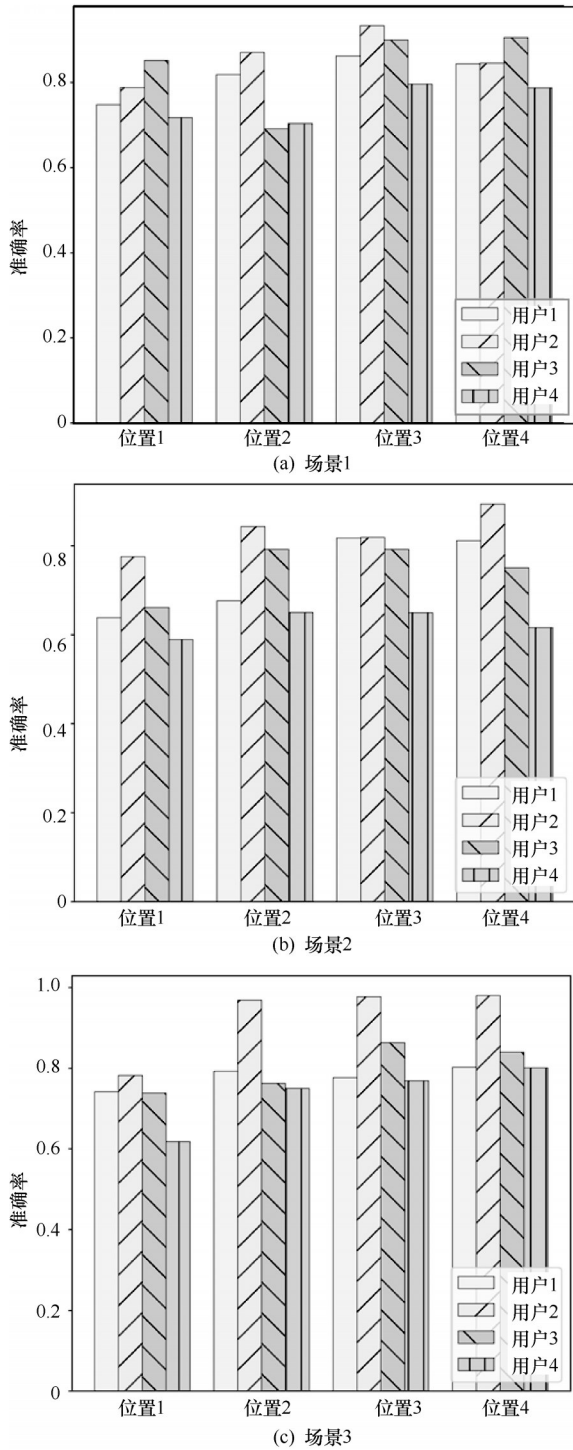


图9 不同用户的跨位置识别性能

图10给出了3个场景下不同位置的跨用户识别性能，图中的横坐标代表目标域。以横坐标用户1为例，其代表了源域的训练数据为用户2、用户3和用户4，预测的目标域为用户2，且依次对每个位置进行跨域感知的实验。同样INDG-Fi模型在不同位置上的实验结果也是有差异的，以场景3下的

跨用户实验结果为例，位置1上的整体表现最优，其跨域准确率分别为0.814 7、0.905 2、0.851 1和0.907 6，而位置3表现最差，其跨域准确率分别为0.640 3、0.801 2、0.767 9和0.824 1。

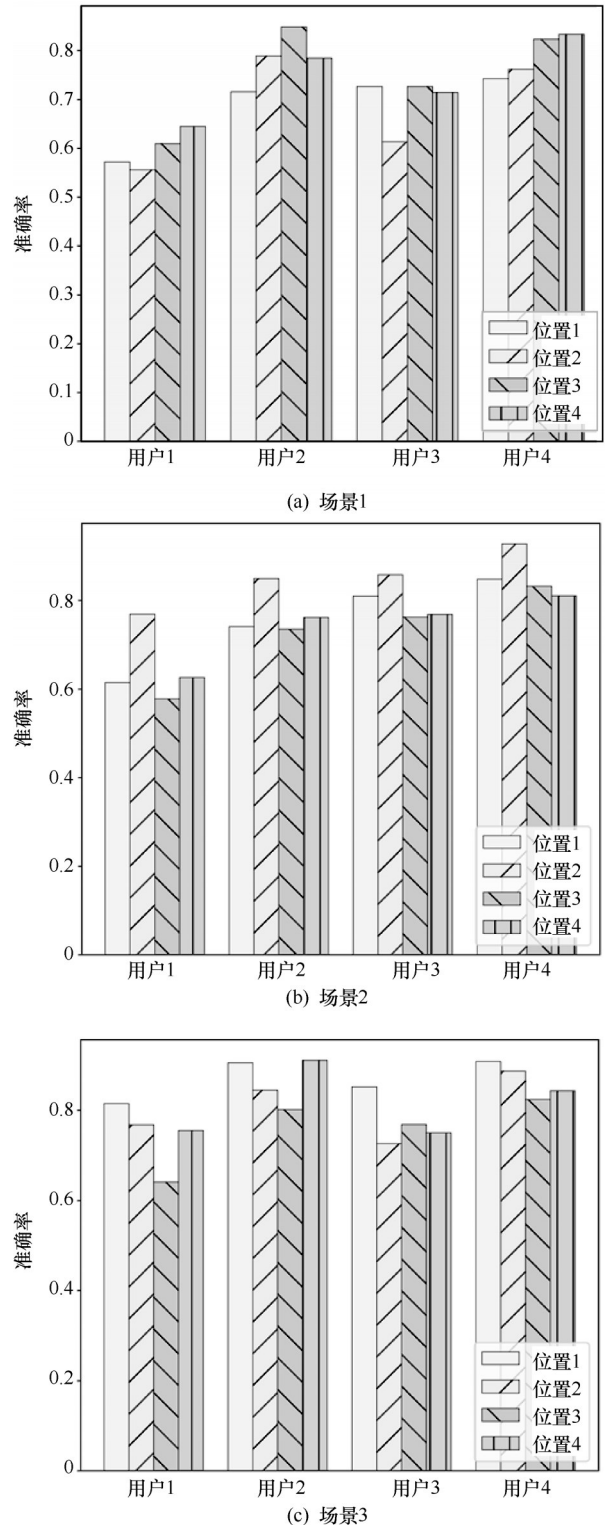


图10 不同位置的跨用户识别性能

### 4.2 特征提取网络比较

本节在编码层使用不同方式的聚合得到序列特征，以此探究不同特征提取网络对模型跨域感知能力的影响。具体而言，本节对比了使用门控循环单元（GRU, gated recurrent unit）、长短期记忆（LSTM, long short-term memory）网络、双向长短期记忆（BiLSTM, bidirectional LSTM）网络以及 Adaptive AvgPool1d 在 3 个场景的跨位置和跨用户的感知性能。GRU 和 LSTM 都是循环神经网络（RNN, recurrent neural network）的变体，可以有效学习序列的时间依赖关系<sup>[24]</sup>。Adaptive AvgPool1d 是一种池化操作，可以适应任意长度的信号输入，并将其平均池化到指定的输出长度，有效地保留输入信号的重要信息<sup>[11]</sup>。实验结果如图 11 所示，可以看出，使用 Adaptive AvgPool1d 整体表现最优，使用 BiLSTM 网络的跨位置感知的最优结果为 0.988 5（用户 1，跨位置的目标域为位置 4），且整体表现不如 LSTM 网络，说明编码层网络使用复杂的时序结构对跨域感知性能提升不明显，复杂的时序结构只能在某些场景下更好地学习到数据分布。

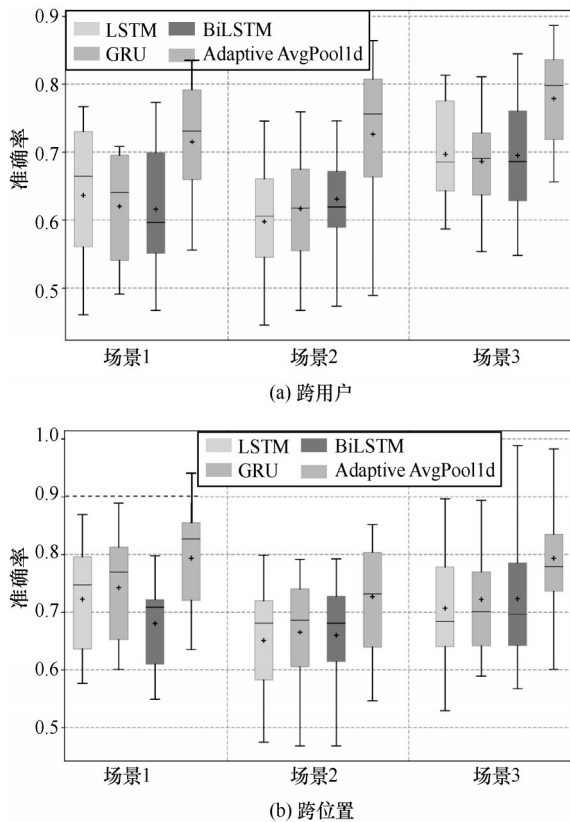


图 11 编码层的聚合方式对跨域感知的影响

### 4.3 与基线模型对比

本节比较 INDG-Fi 模型对比跨域感知的基线模型的性能提升，其中，EI 和 JADA 模型是基于域对抗训练的域自适应算法，WiSR<sup>[11]</sup>和本文 INDG-Fi 模型是基于域泛化学习算法，因此面对目标域不可见时跨域感知表现更好。本文 INDG-Fi 相比 WiSR 优化了编码层网络设计，同时加入了子载波注意力模块，更能帮助模型学习泛化的特征表示，表 2 给出了 3 个场景下跨域感知最高准确率的比较。考虑到本文模型基于 WiSR，图 12 给出了综合的箱线图对比，可以看出 INDG-Fi 模型整体表现更优。

表 2 INDG-Fi 模型与基线模型的最高准确率的比较

场景	类型	EI	JADA	WiSR	INDG-Fi
场景 1	跨用户	0.558 6	0.576 4	0.871 7	0.910 9
	跨位置	0.642 9	0.745 4	0.918 3	0.932 9
场景 2	跨用户	0.589 2	0.766 8	0.900 0	0.927 3
	跨位置	0.662 0	0.727 5	0.896 6	0.893 7
场景 3	跨用户	0.601 9	0.744 0	0.903 4	0.910 9
	跨位置	0.689 8	0.632 1	0.982 8	0.979 9

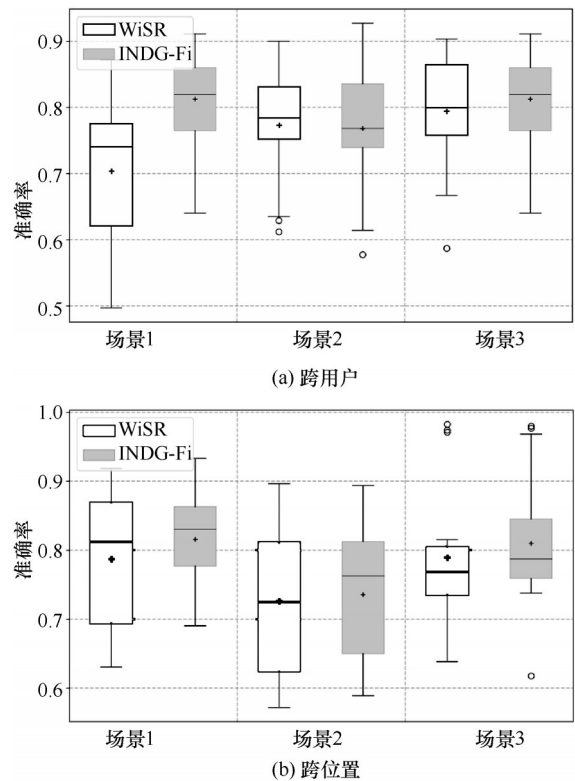


图 12 与基线模型的跨域感知能力对比

### 4.4 接收端数量的影响

上述实验结果均基于一个发射端和一个接收端的数据，而本文数据采集时已经是 2 个接收端同时收集，因此这里实验探究接收端数量对实验结果的影响。表 3 给出了 INDG-Fi 模型使用 2 个接收端的跨域感知能力，其平均准确率在 0.778 3 到 0.855 2 之间，最高准确率在 0.957 7 到 0.993 0 之间，最低准确率在 0.557 3 到 0.687 5 之间。对比使用一个接收端，使用 2 个接收端在 3 个场景下跨用户的平均准确率提升为 1.55%，跨位置的平均准确率提升为 7.66%。

表 3 INDG-Fi 模型使用 2 个接收端的实验结果

场景	类型	平均准确率	最高准确率	最低准确率
场景 1	跨用户	0.778 3	0.982 5	0.600 0
	跨位置	0.855 2	0.957 7	0.687 5
场景 2	跨用户	0.781 1	0.966 1	0.587 2
	跨位置	0.807 2	0.989 3	0.661 0
场景 3	跨用户	0.756 5	0.982 8	0.557 3
	跨位置	0.836 9	0.993 0	0.648 1

为更为直观地展示接收端数量对模型跨域感知能力的影响，图 13 给出了接收端数量对模型跨域感知的影响结果对比。可以看出，2 个接收端整体的实验结果是更优的，这是因为多接收端提供更丰富的空间信息。但这并不意味着 2 个接收端的结果一定优于一个接收端，多接收端可能会面临更多的干扰和信号衰减，导致结果的准确率下降。

### 4.5 公开数据集实验

本节通过在公开数据集 Widar3 上的跨域实验，进一步证明 INDG-Fi 模型对比基线模型 WiSR 的性能提升。选用数据集中“20181109”目录下的数据，该数据集包含 3 名用户在 5 个位置以及 5 个方向采集的 6 种动作数据，分别进行跨位置、跨方向、跨接收端和跨用户实验，实验使用  $F_1$  值作为模型的评价指标，实验结果如图 14 所示。从图 14 可以看出：1) 跨方向的实验结果最优，INDG-Fi 的  $F_1$  值平均为 0.886，WiSR 的  $F_1$  值平均为 0.853，并且 2 个模型  $F_1$  值最高达到 0.99，这是同一动作在不同方向上的变化所引起的域变动相对较小；2) 跨接收端的实验结果较低，INDG-Fi 和 WiSR 的  $F_1$  值平均分

别为 0.534 和 0.599，这是因为每个接收端距离有较大的间隔距离，因此同一动作在不同接收端收到 CSI 模型的变化差距大；3) 跨用户实验结果最差，因为不同用户的体型以及习惯差异造成 CSI 模式差异大，加上仅收集 3 名用户数据，即源域不够丰富，本文模型和 WiSR 在源域数据不足的情况下对域无关特征提取能力较弱。

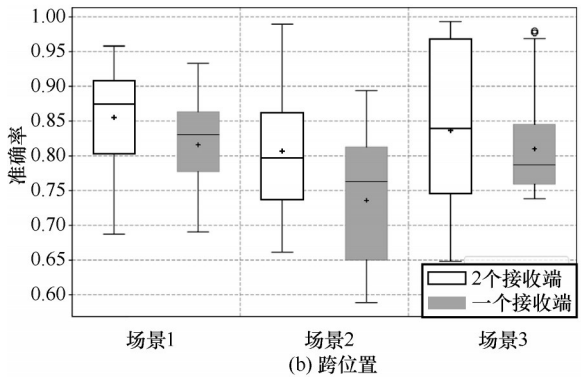
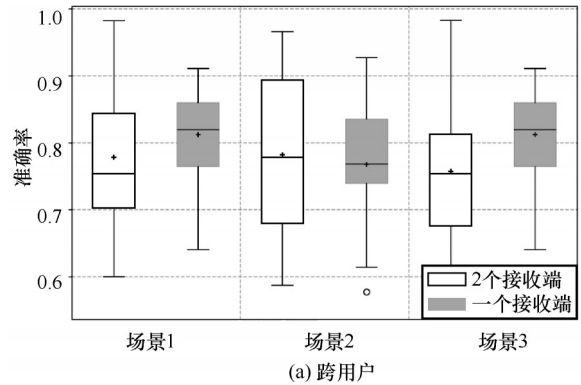


图 13 接收端数量对跨域感知模型的影响

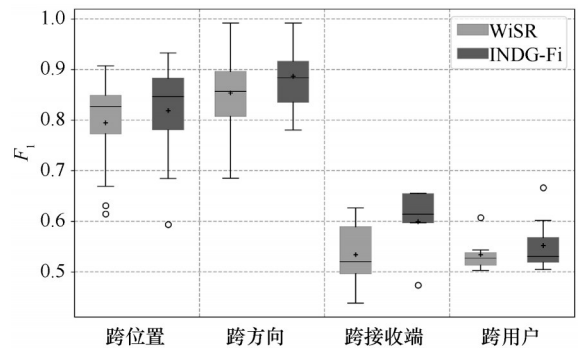


图 14 公开数据集跨域实验结果

### 4.6 消融实验

1) 输入特征消融。探究输入特征对跨域结果的影响，如图 15 所示。采集的 CSI 数据包含幅值和相位，因此输入模型的特征可以为仅幅值、仅相位

和幅值加相位。从图 15 可以看出, 在 3 种情景下, 相位信息对模型影响最显著, 并且幅值和相位信息同时作为输入时跨域的整体结果更优。

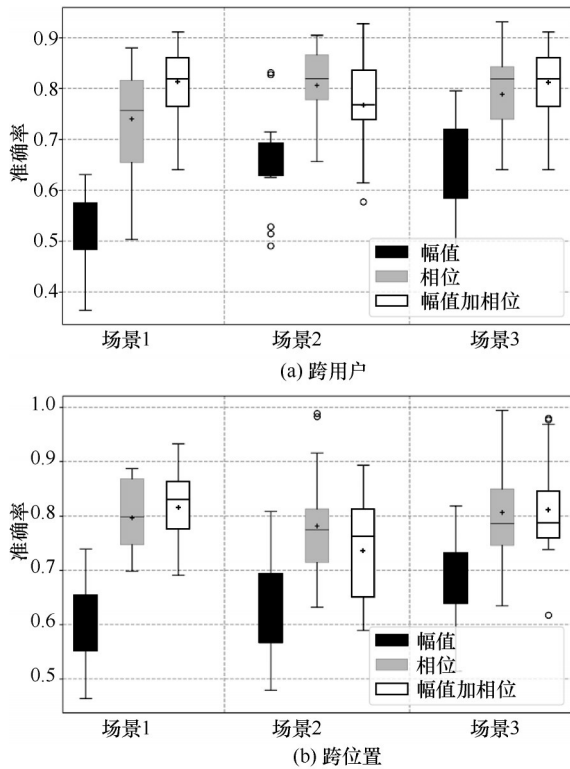


图 15 输入特征消融

2) 数据预处理消融。探究数据预处理对跨域结果的影响, 如图 16 所示。其中, “Without Denoise” 表示输入模型的数据没有经过 Hampel 滤波和线性相位去噪, “With Denoise” 表示数据经过降噪处理。从图 16 可以看出, 对数据进行降噪处理, 可有效提高模型性能。

3) 结构消融。探究模型每个组件对模型性能提升的影响, 如图 17 所示。其中, “without DR and SA” 表示 INDG-Fi 去除掉 DR 模块和子载波注意力模块, “without DR” 表示 INDG-Fi 去除掉 DR 模块, “without SA” 表示 INDG-Fi 去除掉 SA 模块。从图 17 可以看出, 同时添加子载波注意力模块和 DR 模块性能提升最为明显, 并且单独加入 DR 模块或者子载波注意力模块对模型都具有明显的性能提升。

### 5 结束语

针对目标域完全不可见场景下的 Wi-Fi 人体感知问题, 本文提出了一种基于 CSI 实例标准化的域

泛化人体动作识别模型。该模型在编码层中加入子载波注意力模块, 以此关注受人体动作影响更显著的子载波信号, 并通过动作偏向学习和对抗性的域学习将编码层提取的特征偏向仅由人体动作引起的信号特征。由于域泛化学习方式依赖于多个分布不同的源域数据集训练模型, 在某些复杂的动态场景下多源数据集的采集与处理较为耗时耗力。因此, 在未来研究工作中, 将充分考虑在源域数据集受限情形下, 聚焦如何高效学习到可泛化的多维特征表示, 对本文模型进行进一步改进和优化。

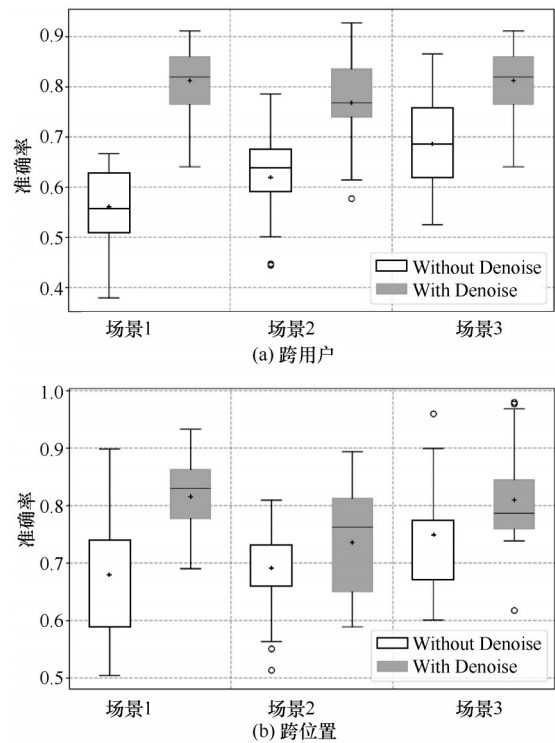


图 16 数据预处理消融

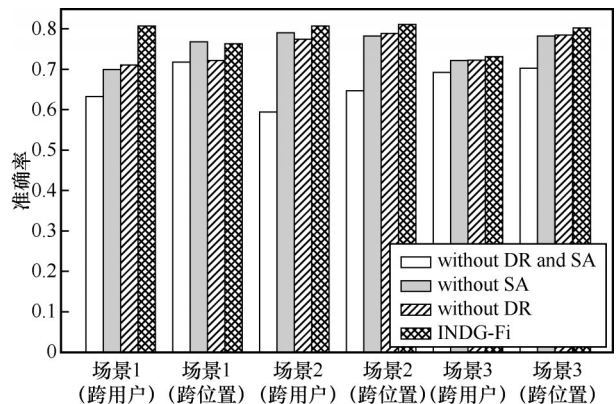
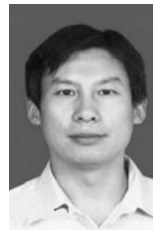


图 17 结构消融

## 参考文献:

- [1] CHEN Z H, ZHANG L, JIANG C Y, et al. WiFi CSI based passive human activity recognition using attention based BLSTM[J]. IEEE Transactions on Mobile Computing, 2019, 18(11): 2714-2724.
- [2] YAO Y C, ZHAO C X, PAN Y H, et al. Human gesture recognition based on CT—a hybrid deep learning model in Wi-Fi environment[J]. IEEE Sensors Journal, 2023, 23(22): 28021-28034.
- [3] TAN S, YANG J, CHEN Y Y. Enabling fine-grained finger gesture recognition on commodity WiFi devices[J]. IEEE Transactions on Mobile Computing, 2022, 21(8): 2789-2802.
- [4] LIU S J, CHEN Y R, WANG H, et al. A low-calculation contactless continuous authentication based on postural transition[J]. IEEE Transactions on Information Forensics and Security, 2022, 17: 3077-3090.
- [5] 周牧, 龙玥辛, 蒲巧林, 等. 基于 CSI 张量分解的室内 Wi-Fi 指纹定位方法[J]. 通信学报, 2021, 42(11): 159-171.  
ZHOU M, LONG Y X, PU Q L, et al. Indoor Wi-Fi fingerprint localization method based on CSI tensor decomposition[J]. Journal on Communications, 2021, 42(11): 159-171.
- [6] WU D, ZENG Y W, GAO R Y, et al. WiTraj: robust indoor motion tracking with WiFi signals[J]. IEEE Transactions on Mobile Computing, 2023, 22(5): 3062-3078.
- [7] YIN G L, ZHANG J Q, SHEN G X, et al. FewSense, towards a scalable and cross-domain Wi-Fi sensing system using few-shot learning[J]. IEEE Transactions on Mobile Computing, 2024, 23(1): 453-468.
- [8] ZHANG X, TANG C P, YIN K, et al. WiFi-based cross-domain gesture recognition via modified prototypical networks[J]. IEEE Internet of Things Journal, 2022, 9(11): 8584-8596.
- [9] ZHANG Y, ZHENG Y, QIAN K, et al. Widar3.0: zero-effort cross-domain gesture recognition with Wi-Fi[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(11): 8671-8688.
- [10] WANG J D, LAN C L, LIU C, et al. Generalizing to unseen domains: a survey on domain generalization[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(8): 8052-8072.
- [11] LIU S J, CHEN Z H, WU M, et al. WiSR: wireless domain generalization based on style randomization[J]. IEEE Transactions on Mobile Computing, 2024, 23(5): 4520-4532.
- [12] GAO R Y, ZHANG M, ZHANG J, et al. Towards position-independent sensing for gesture recognition with Wi-Fi[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2021, 5(2): 61.
- [13] NIU K, ZHANG F S, WANG X Z, et al. Understanding WiFi signal frequency features for position-independent gesture sensing[J]. IEEE Transactions on Mobile Computing, 2022, 21(11): 4156-4171.
- [14] GAO R Y, LI W W, XIE Y X, et al. Towards robust gesture recognition by characterizing the sensing quality of WiFi signals[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2022, 6(1): 11.
- [15] SHENG B Y, XIAO F, SHA L T, et al. Deep spatial-temporal model based cross-scene action recognition using commodity WiFi[J]. IEEE Internet of Things Journal, 2020, 7(4): 3592-3601.
- [16] BU Q R, YANG G, MING X X, et al. Deep transfer learning for gesture recognition with WiFi signals[J]. Personal and Ubiquitous Computing, 2022, 26(3): 543-554.
- [17] JIANG W J, MIAO C L, MA F L, et al. Towards environment independent device free human activity recognition[C]//Proceedings of the 24th Annual International Conference on Mobile Computing and Networking, New York: ACM Press, 2018: 289-304.
- [18] KANG H, ZHANG Q, HUANG Q Y. Context-aware wireless-based cross-domain gesture recognition[J]. IEEE Internet of Things Journal, 2021, 8(17): 13503-13515.
- [19] ZOU H, YANG J F, ZHOU Y X, et al. Joint adversarial domain adaptation for resilient WiFi-enabled device-free gesture recognition[C]//Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). Piscataway: IEEE Press, 2018: 202-207.
- [20] SHI Z G, ZHANG J A, XU R Y, et al. Environment-robust device-free human activity recognition with channel-state-information enhancement and one-shot learning[J]. IEEE Transactions on Mobile Computing, 2022, 21(2): 540-554.
- [21] ZHOU Z P, WANG F, YU J H, et al. Target-oriented semi-supervised domain adaptation for WiFi-based HAR[C]//Proceedings of the IEEE INFOCOM 2022-IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2022: 420-429.
- [22] 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.
- [23] CUI W, ZHANG L, LI B, et al. Received signal strength based indoor positioning using a random vector functional link network[J]. IEEE Transactions on Industrial Informatics, 2018, 14(5): 1846-1855.
- [24] ZHANG Y, QU C, WANG Y J. An indoor positioning method based on CSI by using features optimization mechanism with LSTM[J]. IEEE Sensors Journal, 2020, 20(9): 4868-4878.

## [作者简介]



王杨 (1971-), 男, 安徽灵璧人, 博士, 安徽师范大学教授, 主要研究方向为人工智能、物联网、增强现实。



许佳炜 (1996-), 男, 安徽合肥人, 安徽师范大学硕士生, 主要研究方向为无线感知、表示学习。



王傲 (1999-), 男, 安徽阜阳人, 安徽师范大学硕士生, 主要研究方向为机器学习。



赵传信 (1977-), 男, 安徽凤阳人, 博士, 安徽师范大学教授, 主要研究方向为无线可充电传感器网络、智能信息处理。



夏慧娟 (1997-), 女, 安徽庐江人, 安徽师范大学研究员, 主要研究方向为机器学习、信号处理、计算机视觉等。



季一木 (1978-), 男, 安徽无为, 南京邮电大学教授、博士生导师, 主要研究方向为云计算、大数据、物联网和人工智能等。