文章编号：1003－1251(2026)05－0010－07

数据受限场景下的无设备手势识别方法研究

郭 崇，陶 硕，李鹤群

（沈阳理工大学 信息科学与工程学院，沈阳 110159）

摘 要：针对无设备手势识别(device free gesture recognition，DFGR)技术因环境差异导致的精度下降问题，提出了一种基于深度学习的跨场景手势识别模型 ADAN。该模型的创新之处体现在三个方面：其一，采用改进的 ResNet101 架构构建特征提取器，通过优化卷积核配置，提升捕捉手势细微特征的能力；其二，设计一个基于注意力机制的特征选择模块，通过动态权重分配实现关键特征的过滤和强化；其三，采用改进的域对抗损失函数对不同样本应用不同的权重操作，实现不同环境下的特征对齐。为验证模型性能，本研究基于 Widar 3.0 数据集构建了六个典型的室内环境进行训练。实验结果表明，ADAN 在跨场景手势识别任务中实现了 84.41% 的平均准确率，较传统方法提升了 3.11 个百分点，并优于现有基准模型，为解决基于无线感知手势识别中的场景迁移问题提供了新的技术方法。

关 键 词：无线感知；跨场景手势识别；域对抗损失函数

中图分类号：TP391.41 **文献标志码**：A **DOI**：10.3969/j.issn.1003－1251.2026.05.002

Research on Device-Free Gesture Recognition Method in Data-Limited Scenarios

GUO Chong，TAO Shuo，LI Hequn

（Shenyang Ligong University，Shenyang 110159，China）

Abstract：To address the accuracy decline of device-free gesture recognition (DFGR) technology due to environmental differences，a cross-scenario gesture recognition model based on deep learning，named ADAN，is proposed. The innovation of this model lies in three aspects. Firstly，an improved ResNet101 architecture is adopted to construct the feature extractor，whichi enhances the ability to capture subtle gesture features by optimizing the convolutional kernel configuration. Secondly，a feature selection module based on the attention mechanism is designed，which filters and strengthens key features through dynamic weight allocation. Finally，an improved domain adversarial loss function is used to apply different weight operations to different samples，achieving feature alignment in different environments. To verify the performance of the model，six typical indoor environments for training are constructed in this study based on the Widar 3.0 dataset. The experimental results show that ADAN achieves an average accuracy of 84.415% in cross-scenario gesture recognition tasks，which is 3.115 percentage points higher than that employing traditional method and outperforms existing benchmark models，providing a new technical approach to solving the scene transfer problem in wireless perception-based gesture recognition.

Key words：wireless sensing；cross-scene gesture recognition；domain adversarial loss function

收稿日期：2025－06－27

基金项目：辽宁省教育厅自然科学类项目(JYTMS20230184)

作者简介：郭崇(1980—)，女，副教授，博士；李鹤群(1989—)，通信作者，男，讲师，博士。

随着深度学习技术的蓬勃发展,基于深度神经网络的手势识别方法[1-2]已成为人工智能领域的研究热点。作为一种自然直观的人机交互方式,无设备手势识别(device free gesture recognition,DFGR)凭借无需用户佩戴专用传感器或依赖特定采集设备的特性,极大地提升了交互的便捷性与普适性,使其在普适计算领域展现出重要的应用价值。然而,基于深度学习的 DFGR 方法(deep learning-based DFGR,DL-DFGR)在实际过程中面临严峻挑战,即模型的识别性能高度依赖训练数据对场景的覆盖程度。当训练数据仅来自单一环境时,训练所得的识别模型对环境扰动极为敏感,虽能在训练场景中表现良好,但应用于新环境时识别率会大幅下降。该"场景泛化性差"问题,严重制约了 DL-DFGR 技术的推广,进而推动了跨场景无设备手势识别(cross-scenario DFGR,CS-DFGR)研究的兴起。相较于前者,CS-DFGR 聚焦于提升深度学习模型在未知环境中的适应能力和泛化性能[3-4]。

当前,CS-DFGR 研究主要围绕增强网络可迁移分类能力与优化特征分布对齐两大技术路径展开。在提升网络可迁移分类能力方面,众多学者致力于挖掘与环境无关的手势特征:Shi 等[5]构建的 EI 识别框架,从多个环境收集信号,并从中提炼出与环境无关的特征,显著改善了特定场景下的识别性能;Hu 等[6]借鉴小样本学习思想,设计了 WIGR 手势识别系统。该系统借助特征提取子网与相似性鉴别子网的协同,通过动态调整训练数据分布以模拟真实场景,从而有效增强了模型的领域迁移能力;Wu 等[7]则选取轨迹图像与时间序列作为关键特征,突破环境差异限制;Niu 等[8]基于 Wi-Fi 频率模型量化了信号与目标运动参数的关系,并以手势片段及方向变化特征为核心构建跨场景识别系统。

在利用对抗网络实现特征分布对齐方面,Wang 等[9]采用 Minmax 优化策略,结合中心对齐机制,将目标特征嵌入源特征分布中;Wang 等[10]针对全局对齐的局限性,提出了自适应对抗学习网络,通过局部特征分布校准,有效缓解特征空间偏移问题;Wang 等[11]设计了并行对抗网络,并行提取目标与源场景的公共特征,最终基于相似度度量实现识别;Kang 等[12]提出无链路数量约束的框架,融合对抗学习与特征解纠缠模块,剥离环境干扰信息,为跨场景识别提供了新思路。

针对环境差异引发的 CS-DFGR 性能衰减问题,本文以生成对抗网络(generative adversarial network,GAN)为基础架构,利用 Resnet101-modify[13]构建具备提取细微特征能力的特征提取器,有效解决传统方法特征提取粒度粗糙、细节丢失的弊端。通过引入层次化注意力机制(hybrid adaptive attention module,HAAM)[14]实现特征选择,过滤掉无关特征,并改良域对抗损失函数,实现跨场景特征的精细化对齐,以期提升模型在复杂场景下的手势识别准确率。

1　跨场景无设备手势识别方案

CS-DFGR 方案总体流程如图 1 所示,首先,使用收发设备采集不同手势的信道状态信息(CSI);其次,将采集到的 CSI 数据进行预处理;最后,将处理过的数据输入到本文设计的模型中,经过特征提取、特征选择以及特征对齐实现手势分类。

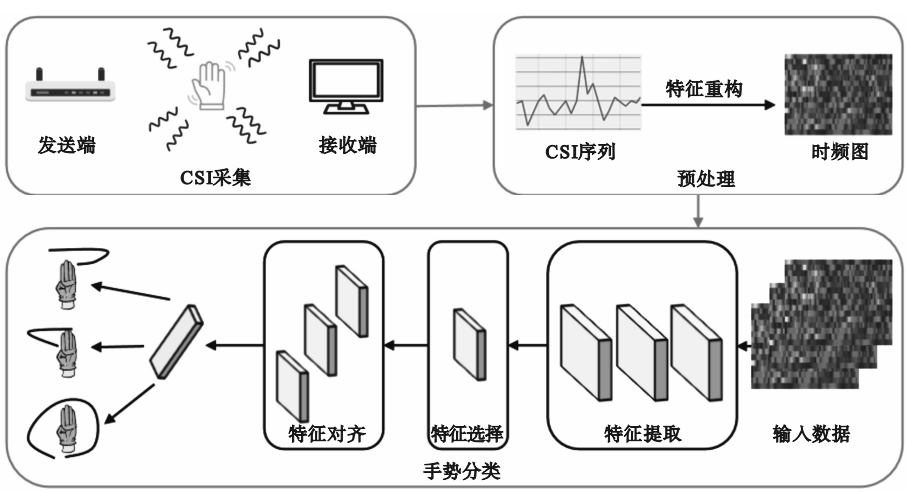


图 1　SC-DFGR 总体流程

Fig. 1　Overall flow chart of CS-DFGR

1.1 信号数据采集

CSI 是在正交频分复用技术中估计信道状态的信息,即每一条无线通道都会有多个子载波,每一条子载波都会有对应的 CSI 值。当无线信号在进行传播时会受到环境的干扰,并体现在接收端的 CSI 数据上。假设接收端信号为 $\boldsymbol{\alpha}$,发射端信号为 $\boldsymbol{\beta}$,则有

$$\boldsymbol{\alpha} = \boldsymbol{H}\boldsymbol{\beta} + \boldsymbol{a} \tag{1}$$

式中:$\boldsymbol{\alpha}$ 为噪声向量;$\boldsymbol{H}$ 表示所有子载波组成的信道矩阵,如果把噪声表示为正态分布的形式,那么就可以估计出信道矩阵 $\boldsymbol{H}$ 为

$$\boldsymbol{H} = \boldsymbol{\alpha}/\boldsymbol{\beta} \tag{2}$$

$\boldsymbol{H}$ 由多个 CSI 值构成,设 CSI 的值为 h,其表达式为

$$h = |x| e^{j\theta} \tag{3}$$

式中:$|x|$ 表示幅度;θ 表示相位偏移。

1.2 数据预处理

采用短时傅里叶变换方法(STFT)将 CSI 数据转换为时频域表示,直观展示信号频率随时间的变化。计算表达式为

$$\boldsymbol{X}(m,\omega) = \sum_{n=0}^{N-1} h(n+m) w(m) e^{-j\omega n} \tag{4}$$

式中:n 是离散时间索引;$h(n+m)$ 表示 $n+m$ 时刻的 CSI 值;$w(m)$ 是窗函数;N 表示所截取信号段包含的样本数;$\boldsymbol{X}(m,\omega)$ 是时间 m 和频率 ω 的二维函数。

STFT 处理得到的数据经可视化映射生成时频图,其样本如图 2 所示。

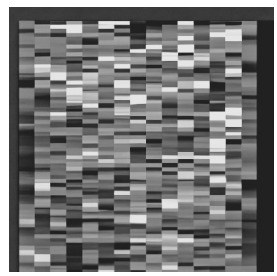


图 2 时频图样本

Fig. 2 Time-frequency diagram

1.3 手势识别模型

手势识别由特征提取、特征选择、特征对齐构成。首先,将上述预处理后的数据输入到 ResNet101-modify 深度卷积神经网络中,通过分组卷积和残差连接协同工作,以提取手势的细微特征;随后,特征选择模块利用注意力机制聚焦于关键特征,并去除数据中的无关噪声;接着,通过对抗机制对来自不同环境的特征进行对齐操作;最后,经由分类器输出分类结果。上述几个模块构成了本文设计的跨场景手势识别模型 ADAN,结构如图 3 所示。

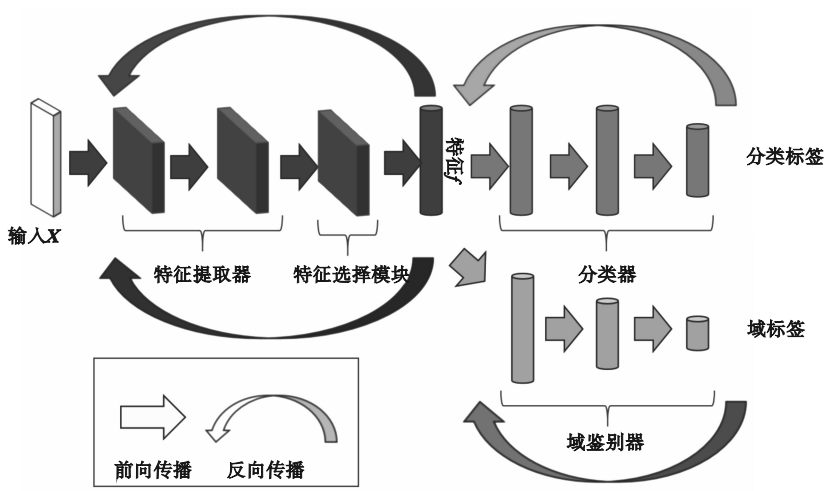


图 3 ADAN 模型结构

Fig. 3 ADAN model structure

2 手势识别模型结构

2.1 特征提取模块

Resnet101-modify 通过构建多层次特征融合机制提取图像中的细微特征,有效增强了模型的特征提取能力,其结构如图 4 所示。

输入先经卷积层、归一化层、ReLU 激活层、池化层完成前期特征提取,再通过包含核心分组残差卷积结构的第一阶段残差层至第四阶段残差层逐步提取更抽象的特征,最终接入 HAAM 模块以完成后续相关操作。

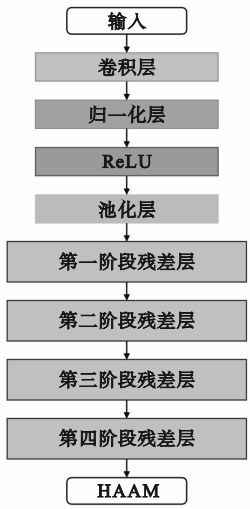


图 4　ResNet101-modify 结构

Fig. 4　ResNet101-modify structure

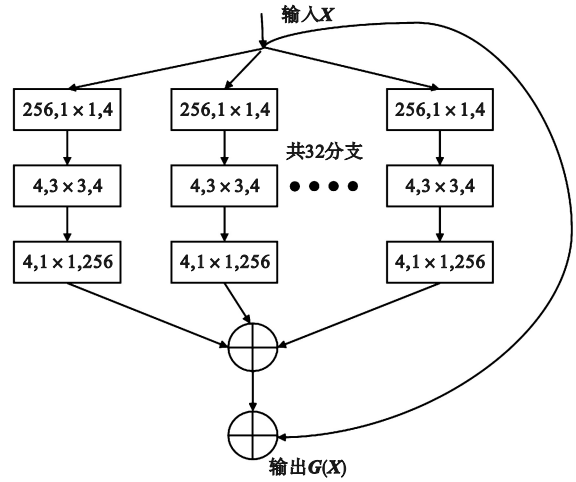


图 5　分组残差模块

Fig. 5　Grouped residual module

结构核心体现于残差层内模块的设计,分组残差模块如图 5 所示。ResNet101-modify 引入基数控制的多分支拓扑结构,结合分组卷积与 Split-Transform-Merge 策略[15],以实现更高效的特征表达。

通过设定基数值直接控制并行分支的数量,可以将模型特征表达能力从单纯依赖深度或通道宽度,扩展到增加独立变换路径数量这一新维度,使模型可以使用更多路径、更窄通道来学习更丰富、更细粒度的特征组合。分组卷积位于模块内部的 3 ×3 卷积层,将输入特征 $\boldsymbol{X}$ 分为 32 个分支并利用 1 ×1 卷积进行降维压缩,压缩后的特征在每个分支上进行独立变换,最后将每个分支得到的变换结果进行聚合并与输入相加得到输出结果,计算表达式为

$$\boldsymbol{G}(\boldsymbol{X}) = \boldsymbol{W}_3 \times \left(\oplus_{i=1}^{32}\left(\boldsymbol{W}_{2,i} \times (\boldsymbol{W}_1 \times \boldsymbol{X})_i\right)\right) + \text{Shortcut}(\boldsymbol{X}) \tag{5}$$

式中:$\boldsymbol{G}(\boldsymbol{X})$ 表示输出特征图;$\boldsymbol{W}_1$ 表示降维投影的 1 ×1 卷积核权重;$\boldsymbol{W}_{2,i}$ 表示分组卷积的第 i 个分支的 3 ×3 卷积核权重;$\oplus_{i=1}^{32}$ 表示通道拼接操作,将 32 个分支的输出沿通道维度拼接,恢复通道数;$\boldsymbol{W}_3$ 表示升维融合的 1 ×1 卷积核权重,是 Resnet-modify 核心参数;Shourcut($\boldsymbol{X}$) 表示残差连接。

2.2　特征选择模块

基于 HAAM 的深度学习模块旨在增强视觉任务特征选择能力,HAAM 注意力模块结构如图 6 所示。

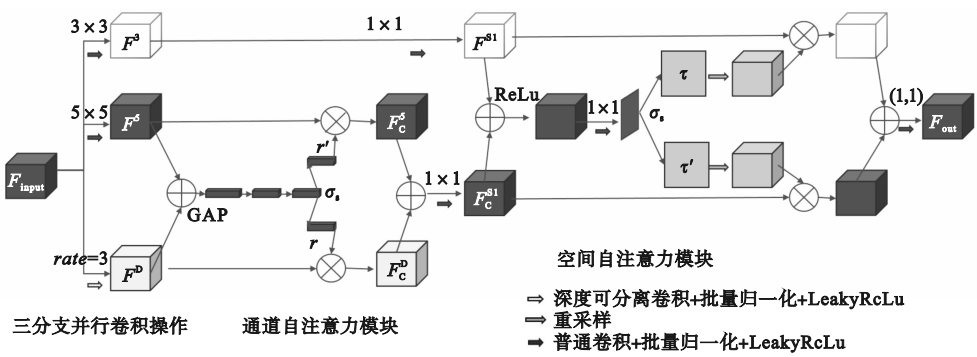


图 6　HAAM 注意力模块

Fig. 6　HAAM attention module

该模块主要包含两个关键部分。第一个部分是三分支并行卷积操作:输入特征被同时输入3 ×3,5 ×5 的普通卷积以及 3 ×3 的深度可分离卷积(以空洞率 *rate* = 3 控制感受野大小)进行处理,原因在于不同场景手势图像分布相似,且边界模糊。3 ×3 卷积可以聚焦局部细节以辅助边界定位,5 ×5 卷积和深度可分离卷积可以捕捉更大范围的上下文信息,减少背景的干扰,通过利用不同

尺寸卷积核的方式可以达到收集多尺度感受野的信息并丰富特征表示的目的。第二个部分是双重注意力优化:首先,提取的多尺度特征被输入通道自注意力模块中,该模块会自动评估每个特征通道的重要性,强化包含重要信息的通道,并去除冗余或干扰性的通道;随后,特征通过空间自注意力模块,经由 1×1 卷积降维至单通道生成空间权重图,聚焦于特征图中最重要的空间区域,提取对识别任务至关重要的局部细节,从而选择最具区分度和可靠性的特征。设输入特征为 $F_{\text{input}} \in \mathrm{R}^{c\times t\times k}$,其中 c,t,k 表示输入特征的通道数、高度、宽度,其输出特征 F_{out} 的计算过程如下。

1)三分支并行卷积:

$$\boldsymbol{F}^3 = \boldsymbol{W}_{3\times 3} \times \boldsymbol{F}_{\text{input}} \tag{6}$$

$$\boldsymbol{F}^5 = \boldsymbol{W}_{5\times 5} \times \boldsymbol{F}_{\text{input}} \tag{7}$$

$$\boldsymbol{F}^{\mathrm{D}} = \boldsymbol{W}^{\mathrm{D}}_{3\times 3} \times \boldsymbol{F}_{\text{input}} \tag{8}$$

式中:$\boldsymbol{W}_{3\times 3}$,$\boldsymbol{W}_{5\times 5}$ 表示 3×3,5×5 卷积核权重;$\boldsymbol{W}^{\mathrm{D}}_{3\times 3}$ 为 3×3 深度可分离卷积核权重。

2)通道自注意力模块:

$$\boldsymbol{F}^{\mathrm{G}} = \mathrm{GAP}(\boldsymbol{F}^5 \oplus \boldsymbol{F}^{\mathrm{D}}) \tag{9}$$

$$\boldsymbol{\gamma} = \sigma_{\mathrm{s}}(\boldsymbol{W}_{\mathrm{fc2}} \times \sigma_{\mathrm{r}})(B(\boldsymbol{W}_{\mathrm{fc1}} \times \boldsymbol{F}^{\mathrm{G}})) \tag{10}$$

$$\boldsymbol{F}^{\mathrm{D}}_{\mathrm{C}} = \boldsymbol{\gamma} \otimes \boldsymbol{F}^{\mathrm{D}} \tag{11}$$

$$\boldsymbol{F}^{5}_{\mathrm{C}} = \boldsymbol{\gamma}' \otimes \boldsymbol{F}^{5} \tag{12}$$

式中:GAP 表示全局平均池化;⊕用于特征逐元素融合;⊗实现注意力权重对特征的校准;$\boldsymbol{W}_{\mathrm{fc1}}$ 表示降维全连接权重;$\boldsymbol{W}_{\mathrm{fc2}}$ 表示升维全连接权重;$B(\cdot)$ 表示批量归一化;$\sigma_{\mathrm{r}}(\cdot)$ 表示 ReLU 激活函数;$\sigma_{\mathrm{s}}(\cdot)$ 表示 Sigmoid 激活函数;$\boldsymbol{\gamma}$ 表示通道注意力权重;$\boldsymbol{\gamma}'$ 是 $\boldsymbol{\gamma}$ 的补集。

3)空间自注意力模块:

$$\boldsymbol{F}^{\mathrm{S1}} = \boldsymbol{W}_{1\times 1} \times \boldsymbol{F}^3 \tag{13}$$

$$\boldsymbol{F}^{\mathrm{S1}}_{\mathrm{C}} = \boldsymbol{W}_{1\times 1} \times \boldsymbol{F}^{5}_{\mathrm{C}} \oplus \boldsymbol{F}^{\mathrm{D}}_{\mathrm{C}}) \tag{14}$$

$$\boldsymbol{\tau} = \sigma_{\mathrm{s}}(\boldsymbol{W}_{1\times 1} \times \sigma_{\mathrm{r}}(\boldsymbol{F}^{\mathrm{S1}} \oplus \boldsymbol{F}^{\mathrm{S1}}_{\mathrm{C}})) \tag{15}$$

$$\boldsymbol{F}_{\text{out}} = \boldsymbol{W}_{1\times 1} \times (\boldsymbol{\tau}' \otimes \boldsymbol{F}^{\mathrm{S1}} \oplus \boldsymbol{\tau} \otimes \boldsymbol{F}^{\mathrm{S1}}_{\mathrm{C}}) \tag{16}$$

式中:$\boldsymbol{W}_{1\times 1}$ 表示 1×1 的卷积核权重;$\boldsymbol{\tau}$ 表示空间注意力权重;$\boldsymbol{\tau}'$ 是 $\boldsymbol{\tau}$ 补集;$\boldsymbol{F}^{\mathrm{S1}}$ 表示空间自注意力初始特征;$\boldsymbol{F}^{\mathrm{S1}}_{\mathrm{C}}$ 表示通道校准后的空间特征。

2.3 特征对齐模块

域对抗学习是一种有效的特征对齐技术。其核心思想是引入域判别器以区分样本属于源域还是目标域;与此同时,特征提取器被训练以生成能够“欺骗”域判别器的特征。通过对抗博弈,让源域和目标域的特征分布尽可能接近。然而,传统的域对抗损失函数存在局限性,其在优化过程中赋予所有样本相同的权重,忽略了不同样本之间可能存在的置信度差异。

熵是衡量样本不确定性的一个有效指标,通过对熵加权操作与域对抗学习结合可以引导模型更有效地进行特征对齐,即在计算域对抗损失时,根据样本的熵值为每个样本分配不同的权重。对于熵值小的样本,赋予较大的权重,使得模型更加关注这些可靠样本对特征对齐的影响;对于熵值大的样本,赋予较小的权重,减少其对特征对齐过程的干扰。重新构造的损失函数 $\mathrm{L}_{\text{adv-ent}}$ 表达式如下。

$$L_{\text{adv-ent}} = -\frac{1}{N_{\mathrm{s}}}\sum_{i=1}^{N_{\mathrm{s}}} w_{\mathrm{s}}^{(i)} \log D(\boldsymbol{f}_{\mathrm{s}}^{(i)}) - \frac{1}{N_{\mathrm{t}}}\sum_{j=1}^{N_{\mathrm{t}}} w_{\mathrm{t}}^{(j)} \log(1 - D(\boldsymbol{f}_{t}^{(j)})) \tag{17}$$

其中

$$w_{\mathrm{s}}^{(i)} = -\sum_{l=1}^{L} p(y_l \mid \boldsymbol{f}_{\mathrm{s}}^{(i)}) \log p(y_l \mid \boldsymbol{f}_{\mathrm{s}}^{(i)})) \tag{18}$$

$$w_{\mathrm{t}}^{(j)} = -\sum_{l=1}^{L} p(y_l \mid \boldsymbol{f}_{\mathrm{t}}^{(j)})) \log p(y_l \mid \boldsymbol{f}_{t}^{(j)})) \tag{19}$$

式中:$\boldsymbol{f}_{\mathrm{t}}^{(j)}$ 表示目标域第 j 个样本的特征向量;N_{s},N_{t} 表示源域和目标域的批量样本数量;$\mathrm{w}_{s}^{(i)}$,$\mathrm{w}_{t}^{(j)}$ 表示源域和目标域样本的熵权重;$p(\mathrm{y}_l \mid \boldsymbol{f})$ 表示分类器将特征 $\boldsymbol{f}$ 预测为类别 l 的概率;L 表示类别总数;$D(\cdot)$ 表示域判别器,输出样本属于源域的概率;$\boldsymbol{f}_{\mathrm{s}}^{(i)}$ 表示源域第 i 个样本的特征向量。

本文通过在域对抗损失函数中增加熵加权损失约束,促使算法在训练过程中更加关注较难分辨的样本,源域和目标域数据达到更好的特征对齐效果,提高模型在跨域任务中的适应能力。

3 实验方案

3.1 数据集描述

本文使用公开数据集 Widar3.0[16] 评估模型性能,其主要在教室、大厅、办公室 3 个地点采集数据。使用的收发设备是配有 Intel 5300 无线网卡的笔记本计算机。每个采集环境的设备和实验布置相同,有一个发射器(每个发射器配备 1 条发射天线)、6 个接收器(每个接收器配备 3 条接收天线),每个接收器每次接收含 90(1×3×30)个子载波的 CSI 数据。本文选取前两个实验地点中收集的六种手势进行实验,分别是平滑,向下,拍手,向前推,画圆,画 Z。其中教室有两个实

验人员,大厅有三个实验人员,每个实验人员的每种手势动作各整理出 750 条,一共是 22 500 条实验数据。并使用该数据构建 6 个跨场景数据子集,分别是教室和大厅环境下的跨位置,跨方向,跨人物。

实验参数设置:本文实验在 RTX3090 显卡上基于深度学习框架 PyTorch 实现。优化器使用随机梯度下降算法(SGD),动量因子为 0.9,正则化参数设置为 0.001,批次设置为 16,初始学习率为 0.01,并使用余弦学习率衰减策略优化学习率,规定最终学习率下降为 0.000 1,根据实验性质将每个数据子集划分为源域和目标域,并使用源域数据以及一部分目标域数据作为训练集,剩余的目标域数据作为测试集用以评估模型效果。

3.2　实验结果

本文采用多场景下取得的平均准确率、精确率、召回率、F1 分数作为评价指标。

3.2.1　对比实验 1

表 1 为基线模型 Resnet101-modify 与 4 种不同的注意力机制组合在 6 种跨场景模式下的对比,其中 C,M,L,H 分别表示与 CBAM[17],Mixed Attention[18],LFGA[19],HAAM 注意力组合。

表 1　不同注意力做特征选择模块的评价结果对比

Table 1　Comparison of evaluation results of feature selection modules with different attention mechanisms.

模型	准确率/%	精确率/%	召回率/%	F1 分数/%
C	83.37	83.13	82.81	82.51
M	79.3	78.48	77.45	77.15
L	81.7	80.76	81.75	81.03
H	84.1	83.48	83.90	83.68

由表 1 可以看出,引入 HAAM 做特征选择模块的模型在综合评估指标上表现最优,其准确率达 84.1%,较基线 ResNet101-modify 提升 1.6 个百分点,在精确率(83.48%)与召回率(83.9%)的提升上也表现明显,判别能力显著。实验进一步表明:注意力机制的有效性极大依赖于主干网络的适配,HAAM 通过分层特征融合实现了更稳定的性能提升。

3.2.2　对比实验 2

本文采用原模型方法和改进后的方法在 6 种跨场景模式下进行对比实验。为验证所提方法的有效性,对基线模型(source)、与 HAAM 注意力机制组合得到的模型(Ours)及对损失函数精准调优得到的最终优化模型(Ours-final)进行性能对比,结果如表 2 所示。

表 2　原模型与改进模型评价结果对比

Table 2　Comparison of evaluation results between the original model and the improved models

模型	准确率/%	精确率/%	召回率/%	F1 分数/%
source	73.58	73.16	72.76	72.58
Ours	84.10	83.48	83.90	83.68
Ours-final	84.41	84.58	84.14	84.35

在基础性能方面,source 模型的各项指标处于中等水平,意味着传统方法在目标任务中仍有改进余地。相较之下,Ours 方法显著提升了模型性能,四项指标较 source 模型分别提升 10.52、10.32、11.14 和 11.10 个百分点,体现出 Ours 方法在特征提取与分类决策上的有效性。Ours-final 模型在保持良好精度的同时,实现了性能的提升,与 Ours 方法相比,提升幅度分别为 0.31、1.10、0.24 和 0.67 个百分点,进一步证实了模型优化策略的有效性。

为验证本文最终方法的跨场景手势识别能力,在相同数据集划分规则下,与 4 种代表性跨场景手势识别模型 EI[5]、ARIL[20]、WIGR[21]、SignFI[22] 进行对比,实验结果如表 3 所示。

表 3　不同跨场景手势识别方法评价结果对比

Table 3　Comparison of evaluation results of different cross-scenario gesture recognition methods

模型	准确率/%	精确率/%	召回率/%	F1 分数/%
EI	76.0	72.90	72.90	72.40
SignFi	77.20	77.30	75.60	75.70
ARIL	71.10	64.70	69.80	65.70
WIGR	81.30	79.10	79.60	79.22
Ours-final	84.41	84.58	84.14	84.35

本文方法四项指标较次优方法 WIGR 分别提升 3.11、5.48、4.54、5.13 个百分点,其中在跨位置、跨人物、跨方向中识别成功样本分别可达 85.4%、85.3%、82.5%,而次优的 WIGR 均只能达到 81% 左右,说明本方法在跨位置、跨用户、跨方向场景的 4 个评价指标上全面超越现有模型,在跨位置场景中,这些模型会将不同测试位置的环境噪声、设备摆放角度偏差纳入特征提取;在跨人物场景中,会将个体手掌大小差异、动作幅度偏差纳入特征提取;在跨方向场景中,无法区分手势朝向变化,导致模型学习到大量冗余信息。而本

文方法通过对特征进一步过滤筛选与领域对齐设计,在保证准确率的同时,能有效解决上述问题,显著降低了对环境稳定性的依赖。

3.2.3 对比试验 **3**

选取现有模型中表现较好的 WIGR 与本文模型 Ours-final 进行可视化对比,观察不同场景特征对齐程度,特征分布图如图 7 所示。相比 WIGR 方法中不同手势类别存在部分重叠、分离效果不明显,本文方法通过特征过滤与领域对齐,在低维空间形成了更紧密的类内聚类和更清晰的类间分离。

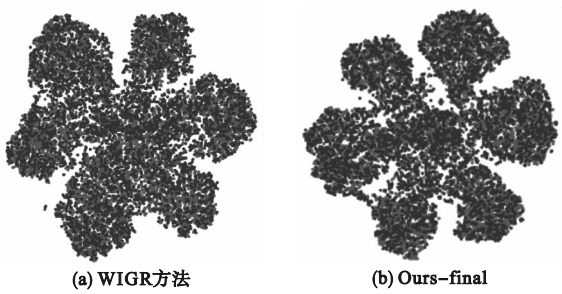


图 7 特征分布图对比

Fig. 7 Comparison of feature distribution maps

4 结论

本文针对无线传感领域中跨场景手势识别问题,提出了基于深度学习的无设备跨场景手势识别模型 ADAN,通过构建多尺度特征提取器联合特征选择机制并使用域对抗机制进行训练,有效改善了由场景因素导致的特征偏移问题,较现有方法准确率提升了 3.11 个百分点。相较于传统方法本研究创新性主要体现在三个方面:在特征提取层面使用 Resnet-modify 提取细微特征;在模型构建方面提出特征选择模块,通过引入层次化注意力机制过滤无关特征;在训练过程中引入熵加权对域对抗损失进行约束,进一步拉近不同场景间的特征距离。

参考文献(References):

[1] 何育浪,赵志彪,李振,等. 基于 DSC-SGRU 模型的 Wi-Fi 手势识别系统研究[J]. 电子测量与仪器学报,2024,38(10):97 – 108.
He Y L,Zhao Z B,Li Z,et al. Research on Wi-Fi gesture recognition system based on DSC-SGRU model[J]. Journal of Electronic Measurement and Instrumentation,2024,38(10):97 – 108. (in Chinese)

[2] Li L S,Bai W Q,Han C. Multiscale feature fusion for gesture recognition using commodity millimeter-wave radar[J]. Computers,Materials & Continua,2024,81(1):1613 – 1640.

[3] Tan L H,Peng Z,Song Y J,et al. Unsupervised domain adaptation method based on relative entropy regularization and measure propagation[J]. Entropy,2025,27(4):426.

[4] 郝海燕,李芳. 基于改进深度子域适应网络的图像分类方法[J]. 沈阳理工大学学报,2024,43(1):69 – 74,90.
Hao H Y,Li F. Image classification method based on improved deep subdomain adaptation network[J]. Journal of Shenyang Ligong University,2024,43(1):69 – 74,90. (in Chinese)

[5] Shi Z G,Zhang J A,Xu R,et al. Towards environment-independent human activity recognition using deep learning and enhanced CSI[C]//GLOBECOM 2020 – 2020 IEEE Global Communications Conference. Taipei,China:IEEE,2020:1 – 6.

[6] Hu P L,Tang C P,Yin K,et al. WiGR:a practical Wi-Fi-based gesture recognition system with a lightweight few-shot network[J]. Applied Sciences,2021,11(8):3329.

[7] Wu J M,Wang J,Dai T,et al. Trajectory features-based robust device-free gesture recognition using mmWave signals[J]. IEEE Internet of Things Journal,2024,11(10):18123 – 18135.

[8] 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.

[9] Wang J,Zhao Y N,Ma X R,et al. Cross-scenario device-free activity recognition based on deep adversarial networks[J]. IEEE Transactions on Vehicular Technology,2020,69(5):5416 – 5425.

[10] Wang J,Wang C C,Yin D Y,et al. Cross-scenario device-free gesture recognition based on self-adaptive adversarial learning[J]. IEEE Internet of Things Journal,2022,9(9):7080 – 7090.

[11] Wang J,Zhao S Z,Lv Y Y,et al. Cross-scenario device-free gesture recognition based on parallel adversarial network[J]. IEEE Transactions on Cognitive Communications and Networking,2024,10(3):893 – 904.

[12] 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.

[13] Xie S N,Girshick R,Dollár P,et al. Aggregated residual transformations for deep neural networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu,HI,USA:IEEE,2017:5987 – 5995.

[14] Chen G P,Li L,Dai Y,et al. AAU-Net:an adaptive attention U-Net for breast lesions segmentation in ultrasound images[J]. IEEE Transactions on Medical Imaging,2023,42(5):1289 – 1300.

[15] Szegedy C,Liu W,Jia Y Q,et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston,MA,USA:IEEE,2015:1 – 9.

[16] 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.

[17] Woo S,Park J,Lee J Y,et al. CBAM:convolutional block attention module[C]//Proceedings of International Conference on Computer Vision(ECCV)2018. Cham,Germany:Springer,2018:3 – 19.

[18] Ashraf M,Chen L H,Zhou X C,et al. A joint architecture of mixed-attention transformer and octave module for hyperspectral image denoising[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2024,17:4331 – 4349.

[19] Shuai C,Zhong J M,Wu S,et al. Locate and verify:a two-stream network for improved deepfake detection[C]//Proceedings of the 31st ACM International Conference on Multimedia. Ottawa,Canada:ACM,2023:7131 – 7142.

(下转第 24 页)