

基于行为多样性数据增强的 RFID 手势识别技术

刘建伟¹, 陈佳彤¹, 姚信威^{2,3}, 韩劲松^{1*}

(1. 浙江大学计算机科学与技术学院, 浙江杭州 310027; 2. 浙江工业大学先进技术研究院, 浙江杭州 310014;
3. 浙江工业大学计算机科学与技术学院, 浙江杭州 310014)

摘要: 手势识别是人机交互中的关键支撑技术, 已广泛应用于智能医疗、虚拟现实和智能家居等场景。相较于依赖视觉信息的方法, 基于射频识别(Radio Frequency Identification, RFID)的手势识别具有非接触、低成本和隐私友好等优势, 展现出良好的应用前景。然而, 现有 RFID 手势识别模型的性能高度依赖于训练数据的规模与多样性, 而射频数据的采集成本较高、数据分布受限, 从而制约了模型在未知样本上的泛化能力。为缓解上述问题, 数据增强被视为提升识别性能的有效手段。然而, 现有数据增强方法往往忽略了人类行为本身的固有差异性, 导致合成样本在分布上的多样性不足。事实上, 即便属于同一手势类别, 不同样本之间仍普遍存在显著差异, 其根本原因在于人类难以以完全一致的方式重复执行同一动作。在手势识别场景中, 这种行为差异在直觉上主要源于手部相对于传感器的距离变化以及运动速度的差异。本文通过建立信号传播模型并结合先验实验, 对这一直觉假设进行了系统验证。基于上述分析, 本文提出一种面向行为多样性的 RFID 数据增强方法。该方法首先从射频信号传播机理出发, 将手势的距离与速度信息显式编码为条件输入。随后, 本文设计了一种新颖的条件扩散模型, 在生成过程中沿距离和速度对应的条件方向进行受控偏移, 从而合成在物理一致性约束下、分布范围更广的高质量样本。通过在手势识别模型的训练集中引入富含行为多样性的合成数据, 可以有效提升模型对未知样本的泛化能力以及手势识别的准确率。本文在真实应用场景中搭建了 RFID 手势识别原型系统, 并开展了系统性的物理实验。实验结果表明, 在 10 类手势识别任务中, 所提出的方法相较于基线模型的识别准确率提升超过 20%, 且显著优于现有的数据增强策略。此外, 该方法的数据增强效果不会随着感知任务规模的扩大而明显衰减。在身份识别任务上的实验结果进一步表明, 本文提出的方法同样适用于其他人机交互应用场景。

关键词: 姿势识别; 射频识别; 数据增强; 多样性; 扩散模型; 无线感知

基金项目: 国家自然科学基金(No.62372400); 中国博士后科学基金(No.2025M781520); 国家资助博士后研究人员计划(No.GZC20241488); 浙江省“尖兵领雁+X”科技计划项目(No.2025C01054, No.2025C01212); 浙江省博士后科研项目择优资助(No.ZJ2025024)

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(XXXX)XX-0001-13

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20251202

Behaviorally Diverse Data Augmentation for RFID-based Gesture Recognition

LIU Jianwei¹, CHEN Jiatong¹, YAO Xinwei^{2,3}, HAN Jinsong^{1*}

(1. College of Computer Science and Technology, Zhejiang University, Hangzhou, Zhejiang 310027, China;

2. Institute of Advanced Technology, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, China;

3. School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, China)

Abstract: Gesture recognition is a fundamental enabling technology for human-computer interaction and has been widely adopted in applications such as smart healthcare, virtual reality, and smart homes. Compared with vision-based approaches, radio frequency identification (RFID)-based gesture recognition offers several distinctive advantages, including contact-free operation, low deployment cost, and privacy preservation, making it a promising solution for practical deployments. However, the performance of existing RFID-based gesture recognition models heavily depends on the scale and diversity of training data. In practice, the acquisition of RF data is costly and the resulting data distributions are inherently limited, which significantly constrains the generalization capability of models to unseen samples. To address this challenge, data augmentation has been widely regarded as an effective means to enhance recognition performance. Nevertheless, most existing augmentation methods overlook the intrinsic variability of human behavior, leading to synthesized samples with insufficient distributional diversity. In fact, even within the same gesture category, substantial variations commonly exist across

different instances, primarily because humans cannot reproduce the same action in a perfectly identical manner. In gesture recognition scenarios, such behavioral variability intuitively arises from changes in the hand's distance relative to the sensor as well as variations in motion speed. This intuitive hypothesis is systematically validated in this work through the establishment of a signal propagation model combined with preliminary experimental analysis. Motivated by these observations, we propose a behavior diversity-aware RFID data augmentation framework. Specifically, grounded in the physical principles of RF signal propagation, the proposed method explicitly encodes gesture-related distance and speed as conditional inputs. A novel conditional diffusion model is then designed to perform controlled shifts along the distance- and speed-conditioned directions during the generation process, thereby synthesizing high-quality samples that are both physically consistent and distributed over a broader behavioral diversity space. By introducing synthetic data enriched with behavioral diversity into the training set of gesture recognition models, the generalization capability to unseen samples and the recognition accuracy can be significantly improved. We implement an RFID based gesture recognition prototype system in real world scenarios and conduct systematic physical experiments. Experimental results show that, in a ten-class gesture recognition task, the proposed method improves recognition accuracy by more than 20% compared with baseline models and significantly outperforms existing data augmentation strategies. In addition, the effectiveness of the proposed data augmentation approach does not degrade noticeably as the scale of the sensing task increases. Experimental results on identity recognition further demonstrate that the proposed method can be effectively applied to other human computer interaction applications.

Keywords: gesture recognition; data augmentation; diversity; diffusion model; wireless sensing

Foundation Item(s): National Natural Science Foundation of China (No.62372400); China Postdoctoral Science Foundation (No.2025M781520); Postdoctoral Fellowship Program of CPSF (No.GZC20241488); "Pioneer" and "Leading Goose" Research and Development Program of Zhejiang (No.2025C01054, No.2025C01212); Postdoctoral Research Excellence Funding Project of Zhejiang Province (No.ZJ2025024)

0 引言

手势识别作为人机交互的核心支撑技术,在智能医疗监护、虚拟现实交互、智能家居控制等领域发挥着重要作用^[1-2]。与传统依赖视觉信息的手势识别方法^[3]相比,基于射频识别(Radio Frequency Identification, RFID)的手势识别技术^[4]具有天然的优势。一方面,其通过射频信号感知人体动作,能够有效规避视觉方法带来的隐私泄露问题;另一方面,该技术具有非侵入性、低成本和无需光照条件等优点^[5-6],因而在智能交互场景中展现出广阔的应用前景。

现有的基于RFID的手势识别方法通常依赖机器学习或深度学习模型^[7-8],将射频信号映射到相应的手势类别。然而,这类模型的性能在很大程度上取决于训练数据的规模及其分布的多样性^[9]。为了在未知样本上获得更优的泛化能力,训练数据需要尽可能涵盖丰富的变化模式。但在实际应用中,大规模采集射频数据往往面临高时间成本和高人力负担^[10]。因此,相较于直接扩展数据采集规模,越来越多的研究采用数据增强方法^[11],通过生成新样本以拓展数据分布,从而提升模型的鲁棒性与泛化能力。

现有针对RFID感知的数据增强方法大致可分为基于数据变换的^[12]和基于生成模型的^[13]两类。前者通常采用旋转、重排等方式对原始数据进行结构化变形,从而引入新的数据排列。虽然能够一定程度上扩展样本集,但生成的数据本质上仍局限于原始样本的

形态变换,新增分布有限。后者则利用深度生成式神经网络(如生成式对抗网络^[13]和扩散模型^[14]),通过学习数据的内部结构来从噪声中重构出新的样本,从而突破原始样本集的局限。然而,由于生成模型往往只能学习训练数据中已有的分布,且更重要的是,忽略了人体行为固有的多变性,导致生成的数据多样性受限,数据增强效果受限。

事实上,已有研究^[15]表明,人类难以以完全一致的方式重复执行同一个动作。直观来看,在RFID手势识别场景中,人手相对于RFID标签的距离变化以及人手运动速度的变化,是导致同一手势不同样本之间产生显著差异的关键因素。若样本生成和数据增强过程未显式考虑距离与速度的多样性,训练数据中仅覆盖有限的距离与速度分布范围,则所训练的手势识别模型在面对未知样本时将难以具备良好的泛化能力。为验证上述直觉假设,本文针对RFID手势识别场景建立了射频信号传播模型,系统分析了距离与速度变化对信号特征分布的影响,结果表明二者是类内信号样本多样性的重要来源。同时,本文通过先验实验进一步证明,在识别模型的训练数据中引入距离与速度这两类行为因素,能够有效提升手势识别性能。

基于上述分析,本文提出一种面向行为多样性的RFID数据生成方法,通过在生成过程中显式引入距离多样性与速度多样性等行为变化因子,有效扩展生成数据的分布空间。具体而言,本文首先构建射频信

号传播模型,将样本之间的距离变化与速度变化统一量化为偏移编码。该编码在保留真实信号传播物理特性的同时,能够刻画样本之间随时间演化的行为差异。进一步地,本文设计了一种面向RFID数据变换的扩散生成模型。不同于现有从随机噪声出发直接生成虚拟样本的无线数据生成方法^[13-14],本文所提出的扩散模型以距离与速度偏移编码为条件,引导模型学习将一个样本沿特定行为变化方向逐步转换为另一种状态的过程。在实际生成过程中,本文通过对部分已知手势类别的真实样本进行两两配对,计算大量距离与速度偏移编码,并利用这些编码对新类别样本进行受控转换,从而合成大量具有物理一致性且行为多样的虚拟样本。该方法能够显著丰富有限训练数据的行为分布,有效降低真实数据的采集开销。

为了验证所提方法的有效性,本文构建了真实场景实验平台,并邀请5名志愿者参与实验,覆盖10类常见手势。实验结果表明,采用所提出的生成方法进行数据增强后训练得到的模型,在不同手势识别任务中的准确率相较于仅使用原始数据训练的基线模型提升超过20%,显著优于现有的面向RFID感知的数据增强策略。此外,提出的数据增强方案可被扩展至其他感知应用中,例如基于RFID的身份识别。

本文的主要贡献包括三个方面:

(1) 本文针对基于RFID的手势识别场景进行了物理建模,理论与实验验证表明距离与速度是引起同类样本间差异的关键因素;

(2) 本文提出了一种面向行为多样性的RFID数据增强方法,有效提升了有限数据的可用性,并可通过高保真、多样化的合成数据替代部分真实数据用于识别模型训练,从而降低用户的数据采集开销;

(3) 本文在真实环境中通过多名志愿者开展了识别实验。实验结果表明提出的数据增强方法能有效提升手势识别性能,且可有效扩展到身份识别等更多感知应用中。

1 相关工作

本研究主要与两类研究相关,即姿势识别与无线感知数据增强。本章将简要概述这两类工作,并总结本研究与先前工作的区别。

1.1 姿势识别

姿势识别(如手语识别和手势识别)是众多人机交互应用的核心支撑技术。传统的姿势识别方法通常依赖摄像头^[16-17]、可穿戴设备^[18-19]或声纳^[20-21]来获取人体动态信息,但这些方法均存在一定局限性。例如,基于视觉的方案不可避免地涉及用户隐私泄露问题;可穿戴设备容易造成佩戴不适,难以长期使用;

而声纳则易受环境噪声干扰。近年来,研究者开始探索利用RFID来实现非接触式姿势识别的可能性。与传统方法相比,基于RFID的姿势识别具备保护视觉隐私、非侵入性以及鲁棒性更强等优势^[12],因而逐渐成为重要的研究方向。

现有的基于RFID的姿势识别方法通常依赖机器学习或深度学习模型,将信号指示量(如信号强度或相位信息)映射为具体的姿势类别。例如,Dian等人^[22]提出了一种多标签数据采集装置,通过获取人体运动的时空特征,再结合深度学习模型有效消除环境干扰,从而实现稳定的手势识别。为实现对手势的分类,Zhang等人^[12]则从信号处理角度出发,对原始数据进行时域和结构上的变换,并利用卷积神经网络(Convolutional Neural Network, CNN)将多维特征映射为手势类别。Ma等人^[23]则关注数据采集效率问题,提出了一种基于信号置换的优化策略,并结合神经网络提升了识别性能。另一项由Zhang等人^[7]主导的研究从相位信息中提取统计和小波变换特征,再分别使用随机森林、朴素贝叶斯分类器以及卷积神经网络进行手势识别,验证了多模型识别的可行性。与此同时,Wang等人^[8]设计了专门的多标签采集平台,通过理论建模提取细粒度射频特征,并将其转化为图像形式,最终利用CNN实现了手指追踪与手势分类。

尽管基于学习的方法在识别精度上取得了显著进展,但其性能高度依赖于训练数据的数量与多样性。鉴于大规模射频数据采集在时间和人力成本上的高昂代价,研究者们逐渐转向数据增强,通过生成新的样本来扩展数据分布,从而提升模型的泛化能力。本文提出了一种新颖的基于行为多样性的RFID数据生成方法,通过在数据合成过程中显式地引入丰富的距离和速度变化,有效提升了合成数据的可用性和手势识别的准确率。

1.2 无线感知数据增强

数据增强作为基于学习的感知方法中的重要环节,在提升模型的泛化能力和鲁棒性方面发挥了关键作用。通过对已有数据进行变换或合成,数据增强不仅能够有效缓解训练样本不足的问题,还能让模型在面对复杂和多变的环境时保持稳定的性能。此外,数据增强还能避免过拟合,使模型学习到更具普适性的特征。因此,它已成为无线感知^[24]、计算机视觉^[25]以及语音处理^[26]等多个领域中不可或缺的技术手段。

在RFID感知任务中,研究人员同样引入数据增强以提升感知性能。现有方法大致可分为基于数据变换的和数据驱动的两类。前者通常通过维度交换对原始数据进行直接变形,从而在原始分布的基础上生成新的数据。例如,Zhang等人^[12]在利用标签阵列

采集信号后,将数据在阵列所在的二维平面上分别旋转 90° 、 180° 和 270° 以生成新样本,并结合时间反转与跨样本拼接来扩展训练集。Ma等人^[23]则通过对不同类别样本进行组合,构造新的训练数据。然而,这类方法通常仅在既有数据分布上进行变换,所生成样本的多样性较为有限;同时,其往往忽略射频信号在物理世界中的真实传播机理,因而难以保证合成样本的物理一致性与信号保真度。相比之下,数据驱动的方法通常借助深度生成式模型进行数据合成,能够更充分地刻画射频信号内部的复杂结构特征,从而有效提升合成数据的保真度。例如,Yang等人^[13]利用循环神经网络构建生成器,并通过全连接层设计判别器,结合对抗学习的思想不断优化生成与判别过程,使得判别器具备更强的区分能力,而生成器则能够合成高保真的RFID数据。近年来,研究人员提出了扩散模型^[27]进行数据生成。相比于生成式对抗网络,扩散模型在训练过程中更加稳定,不易出现模式崩溃问题,并能够更全面地覆盖数据分布,从而在生成质量和多样性方面展现出显著优势。例如,Wang等人^[14]利用扩散模型作为基础的生成框架,通过输入动作的类别标签为条件,约束模型的去噪过程,生成高保真的射频信号表达。

然而,现有面向无线感知的数据生成方法普遍忽略了人类行为本身所具有的固有多样性,导致生成过程缺乏明确的条件引导,进而限制了合成样本在行为空间中的覆盖能力。针对这一问题,本文提出了一种面向行为多样性的生成方法,在数据合成过程中显式引入距离变化与速度变化作为行为约束,对生成方向进行有效引导。在保证数据物理一致性与信号保真度的同时,该方法显著提升了合成数据的行为多样性,从而有效提高了有限真实数据的利用效率,并增强了手势识别模型对未知样本的普适性与识别性能。

2 研究动机

在RFID^[28-29]感知中,常用的信号表征量包括接收信号强度、幅值以及相位等。其中,相位信息虽然能够提供较为精细的传播特征,但往往容易受到硬件漂移和环境噪声的干扰;幅值则需要额外的解调处理,稳定性相对有限。相比之下,信号强度具有获取简便、鲁棒性较强以及硬件实现成本低等优势。因此,本文选取信号强度作为表征人体运动特征的主要指标。为验证在手势识别场景中,“人手相对于RFID标签的距离以及人手运动速度是导致同类信号样本之间差异的关键因素”这一直觉假设,本章首先构建了如图1所示的信号传播模型,并通过理论分析对该假设的合理性进行了系统论证。随后,结合仿真实验

进一步验证了理论分析的有效性,同时揭示了在识别模型的训练数据中引入距离与速度多样性对于提升识别性能的必要性。

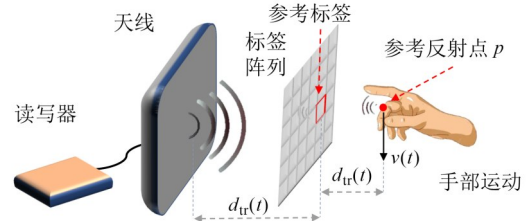


图1 基于RFID的手势识别场景

Figure 1 Scenario of RFID-based gesture recognition

2.1 多样性分析

在典型的RFID手势识别场景中,系统由读写器和标签阵列组成。手部在标签阵列后方运动以实现手势录入。读写器向标签发射连续波信号,标签通过反向散射将信号反射回读写器。当传播路径上存在手部时,信号中将叠加其反射特征^[30],最终读写器接收到的信号包含了手部运动引入的动态特征,可用于手势识别^[31]。

为便于分析,以下以单个标签为例建立传播模型。读写器发射信号可表示为

$$s_{tx}(t) = A_0 e^{j2\pi f_c t} \quad (1)$$

其中, A_0 为发射信号幅度; f_c 为载波频率。以手部某一反射点 p 的反射为例,经过标签反向散射后返回读写器的路径可建模如下:设该反射点与标签的瞬时距离为 $d(t)$,其沿传播方向的运动速度为 $v(t)$ 。则手部反射路径引入的相位延迟和多普勒频移可分别表示为^[31]

$$\phi_{hand}(t) = \frac{2\pi}{\lambda} 2d(t) \quad (2)$$

$$f_D(t) = \frac{2v(t)}{\lambda}$$

其中, λ 为信号波长。在单径近似下,传播损耗可用路径损耗模型表示。传播损耗遵循幂律衰减^[32]: $L(d) \propto d^{-\gamma}$,其中, γ 为路径损耗指数。令 σ_{tag} 为标签的反向散射系数, σ_{hand} 为手部反射强度,则读写器从单个标签接收到、包含手部二次反射成分的信号可表示为

$$r(t) = A_0 \cdot \sigma_{tag} \cdot \sigma_{hand} \cdot \frac{1}{d_{tr}(t)^{\gamma_1}} \cdot \frac{1}{d_{tr}(t)^{\gamma_2}} e^{j\left(2\pi f_c t + \phi_{hand}(t) + 2\pi \int_0^t f_D(\tau) d\tau\right)} + n(t) \quad (3)$$

其中, $d_{tr}(t)$ 表示手部反射点到标签的传播距离; $d_{tr}(t)$ 表示标签到读写器的传播距离; γ_1 和 γ_2 为对应链路的路径损耗因子; $n(t)$ 表示环境噪声。若将幅度项合并为 $A(t)$,相位项合并为 $\Theta(t)$,则可得到

$$r(t) = A(t)e^{i\theta(t)} + n(t) \quad (4)$$

当人手在标签前以特定轨迹运动时,可以将参考反射点在时间区间 $[0, T]$ 内的位移近似表示为

$$d(t) = d_0 + s(t) \quad (5)$$

其中, d_0 为初始距离; $s(t)$ 为随时间变化的位移函数,其时间导数即速度: $v(t) = \dot{s}(t)$ 。接着,接收信号的幅度序列可表示为

$$A(t) = K \cdot \frac{1}{(d_0 + s(t))^r} \quad (6)$$

其中, K 为常数(包含发射幅度与反射系数等)。同时,相位序列可表示为

$$\Theta(t) = 2\pi f_c t + \frac{4\pi}{\lambda}(d_0 + s(t)) + \frac{4\pi}{\lambda} \int_0^t v(\tau) d\tau \quad (7)$$

因此,接收信号的强度包络可表示为

$$|r(t)| \approx A(t) \cdot |\cos(\Theta(t))| + n(t) \quad (8)$$

结合式(6)、(7)以及(8)可以看出,信号的强度变化直接受距离和速度的影响。当距离 $d(t)$ 增加时(反射点远离标签),幅值按 d^{-r} 衰减,信号强度显著下降,且在近距离下对距离变化更加敏感;而速度 $v(t)$ 的增加会加快幅度的变化速率,显著影响强度包络的快速波动。因此,即便是相同手势,在不同次录入中,由于距离与速度的差异,采集到的信号也会表现出不同分布。因此,在数据生成中引入距离与速度的变化,是扩展信号多样性的重要途径。

2.2 先验实验

由上述理论分析可知,人手在不同识别尝试中,其距离和速度的变化是导致同一类别的信号样本间呈现多样性的重要来源。为进一步验证这一推导的合理性与有效性,本节在加入随机噪声的条件下设计并开展了两个先验仿真实验。

在第一个实验中,固定手部反射点的移动速度,分别设置三种不同的相对初始距离(D1=0.5、D2=1.0、D3=1.5),并在标签前绘制“1”的轨迹,估算对应的接收信号强度。每种距离下生成100个样本,经UMAP(Uniform Manifold Approximation and Projection)算法^[33]降维后其二维分布如图2所示。结果表明,相同轨迹在不同距离下形成明显分簇,而同一距离下的样本分布相对较近,说明手部距离确实是引起样本分布差异的重要因素。为进一步分析上述多样性对识别性能的影响,本实验在原有实验设置基础上,引入另一种绘制“(”的运动轨迹,并在每一种距离条件下仿真生100个样本。随后,采用逻辑回归(Logistic Regression, LR)分类器^[34]对生成样本进行分类实验。具体而言,当仅从距离为D1的样本中随机选取30条作为训练集,并从所有距离的样本中随机选取100条

作为测试集时,分类准确率为61%。当训练集(仅从D1的样本抽取)的样本数量增加至60条时,识别准确率仅小幅提升至63%,且继续扩大该类样本的规模并未带来进一步的准确率提升。相比之下,当训练样本从三种不同距离的样本中随机抽取时,识别准确率显著提升至82%。上述结果表明,在识别模型训练过程中,若仅增加训练样本数量而未引入距离维度的多样性,其对识别性能的提升作用十分有限;相反,通过引入距离的多样性,可以有效扩展训练数据的分布范围,从而显著提升识别性能。

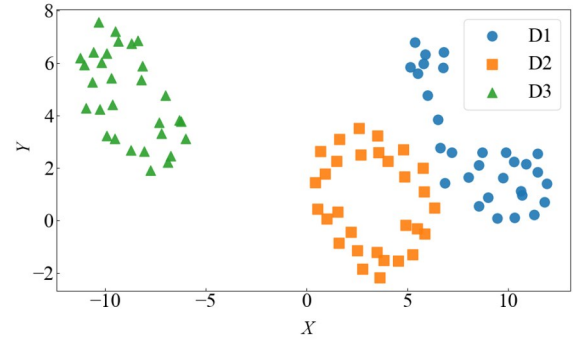


图2 不同距离下信号强度的分布

Figure 2 Signal distributions under different distances

在第二个实验中,固定初始距离,改变反射点的相对速度($V_1=0.5$ 、 $V_2=1.0$ 、 $V_3=1.5$),重复生成100个样本并进行UMAP可视化。图3的结果显示,不同速度的样本分布在不同簇中,而相同速度下的样本则相对紧密聚集。这表明速度同样是样本多样性的关键诱因。与距离相关实验类似,本节同样针对轨迹“(”进行信号仿真并开展分类实验。实验结果表明,当训练数据仅来自单一速度条件时,模型在包含多种速度的混合样本上的测试准确率仅为65%,且即使进一步增加该单一速度条件下的训练数据,准确率也未出现明显提升。相比之下,当训练数据覆盖多种速度条件时,模型的测试准确率可显著提升至86%。这进一步验证了仅通过增加训练样本数量难以有效提升识别性能的结论;相比之下,引入更丰富的速度变化能够显著扩展训练数据分布,从而有效增强分类器对未知样本的普适性与泛化能力。

由上述理论分析与仿真实验可知:距离与速度的差异是导致同类样本间存在多样性的重要因素。若要充分利用有限数据并提升模型对未知样本的泛化能力,增强过程中必须显式引入距离与速度的多样分布。

3 面向行为多样性的数据增强

本章将系统阐述如何基于有限的RFID感知数据

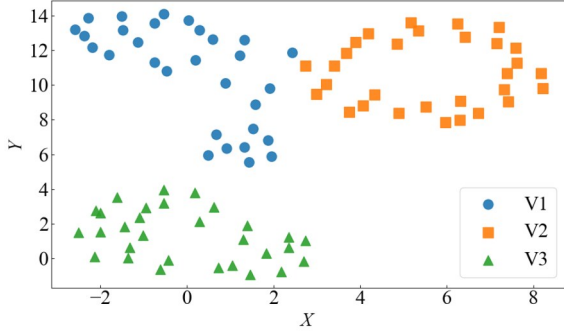


图3 不同速度下信号强度的分布

Figure 3 Signal distributions under different velocities

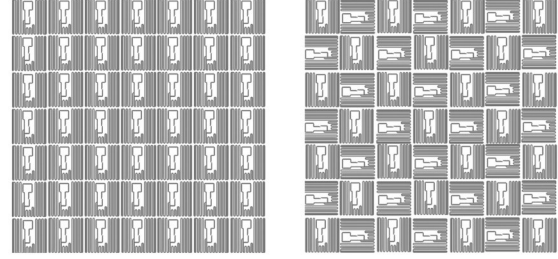
合成大量具备行为多样性的虚拟数据,从而有效提升识别模型的性能。首先,对原始RFID信号进行必要的预处理;随后,基于射频信号传播模型,对样本在距离与速度维度上的变化进行偏移编码;最后,构建条件扩散模型,以距离和速度偏移作为条件约束,生成兼具高保真度与行为多样性的合成数据。

3.1 信号预处理

在RFID感知系统中,标签阵列的几何排布对信号的质量与分辨率有直接影响。若标签数量过少,将无法覆盖手部动作的主要范围,导致采集到的信号难以充分刻画动作的空间特征;若标签数量过多,则会引发严重的电磁耦合效应,部分标签信号强度下降,甚至出现标签无法被正确读取的情况。为在空间分辨率与信号稳定性之间取得平衡,采用 7×7 的二维标签阵列,总计49个标签。这一规模既能够覆盖手部在自然动作过程中的舒适运动空间,也能保证系统在读取性能与精度之间取得合理的折中。在阵列的具体构建中,相邻标签的天线方向并非采用如图4左侧所示的同向排列,而是采取右侧所示的两两垂直排布方式,即相邻的两个标签天线之间方向互相垂直。该设计能够有效减轻标签间的强耦合效应,从而避免信号过弱甚至不可读取的现象。假设阵列坐标为 $\mathcal{T} = \{T_{ij} | i, j \in \{1, 2, \dots, 7\}\}$,其中, T_{ij} 表示位于阵列第*i*行、第*j*列的标签。通过合理的方向配置,每个标签在信号采集中的作用得到充分发挥,从而保证了阵列整体的稳定性与数据的可靠性。下文将阐述如何对该标签阵列采集的数据进行进一步处理以便于后续模型训练。

在RFID感知实验中,并非所有采集的信号均包含手部姿势信息。若对整个连续序列直接采集与存储,将带来大量冗余数据,增加存储开销并削弱模型对关键信息的学习能力。因此,需要设计一种手势检测机制,以在实时采集过程中判定人体动作是否发生。

具体而言,手势检测采用阵列中心位置的标签作为基准。设该中心标签在时刻*t*的信号强度为 $S(t)$,



① 同向标签布置方式

② 垂直标签布置方式

图4 不同的标签阵列布置方式

Figure 4 Different layouts of tag array

在长度为 W 的滑动窗口中,计算其中位数 $\text{median}(S(W))$ 及中位数绝对偏差(Median Absolute Deviation, MAD)^[35]:

$$\text{MAD}(W) = \text{median}(|S(t) - \text{median}(S(W))|) \quad (9)$$

进一步地,计算窗口内每个点的偏差,若存在:

$$|S(t) - \mu_w| > \theta \cdot \text{MAD}(W) \quad (10)$$

则判定在该时间窗口内有动作发生。其中, μ_w 表示窗口 W 内信号的均值; θ 为经验设定的阈值。一旦满足该条件,即视为手部动作的起始时刻,并从该窗口开始对信号进行截取。每个标签在一次动作中采集 N_{time} 个时间点的数据。由此,每个手势样本构成一个二维矩阵 $\mathbf{X} \in \mathbb{R}^{49 \times N_{\text{time}}}$,其中每一行对应一个标签的动态信号序列,每一列表示某一时刻所有标签的信号强度值。该矩阵全面记录了手势执行过程中的空间与时间特征,可被用于后续的手势识别。

接着,由于不同标签采集到的信号强度范围可能存在显著差异,若直接输入模型,容易造成部分标签在特征学习中占据主导,而其他标签信息被弱化。为了解决这一问题,需要在样本层面进行统一的尺度变换。本文采用Min-Max归一化方法^[36],将整个样本矩阵的数值范围线性映射到区间 $[0, 1]$ 。对于样本矩阵 $\mathbf{X} \in \mathbb{R}^{49 \times N_{\text{time}}}$ 以及其每个元素 $X_{i,t}$,首先计算矩阵的全局最小值 X_{\min} 与最大值 X_{\max} :

$$X_{\min} = \min_{i,t} X_{i,t}, X_{\max} = \max_{i,t} X_{i,t} \quad (11)$$

随后,对每一个采样点进行归一化处理:

$$X'_{i,t} = \frac{X_{i,t} - X_{\min}}{X_{\max} - X_{\min}} \quad (12)$$

其中, $X'_{i,t}$ 为归一化后的数值。通过该方法,样本的所有信号强度被映射到 $[0, 1]$ 区间,既保留了不同标签信号的相对差异,又消除了整体幅值的绝对差异。这种归一化方法不仅避免了高强度信号在特征学习过程中占据主导地位,又能够有效缓解梯度消失或梯度爆炸问题,从而加快模型收敛并提升训练的稳定性。

3.2 距离与速度偏移编码

由前文分析可知,在手势执行过程中,信号表现的差异主要受人手与标签阵列之间的距离以及手部运动的速度的影响。这两类行为因素是同一手势类别的样本多样性的主要来源。因此,在数据生成阶段引入距离与速度的多样性,有助于有效扩展合成样本的分布范围,从而提升识别模型对未知样本的泛化能力。

一种直观的方法是在数据生成过程中直接指定距离或速度的标量值。然而,该方式存在两方面不足:其一,任意设定的标量值可能偏离真实的手势执行过程,导致合成数据不符合物理规律;其二,完整的手势属于随时间演化的动态过程,单一标量难以刻画动作过程中距离与速度的连续变化特征。针对上述问题,本文提出一种基于样本对比的行为偏移编码方法。具体而言,从现有训练集中随机选取两个属于同一手势类别的样本,分别记为样本1和样本2。通过对比两者在时序信号上的差异,提取能够表征距离与速度偏移的时序信息。该方法在保留射频信号物理传播特性的同时,有效刻画了贯穿整个手势执行过程的距离和速度变化,为后续数据生成提供了多样的行为约束。

3.2.1 距离偏移编码

在距离变化的编码方面,考虑到信号强度与传播距离之间的反比关系,采用路径损耗模型进行量化。对信号强度矩阵 $\mathbf{X} \in \mathbb{R}^{49 \times N_{\text{time}}}$,其元素 $X_{i,t}$ 表示第 i 个标签在时刻 t 的接收信号强度。根据路径损耗模型^[32],可知:

$$P_r(d) \propto \frac{1}{d^\gamma} \quad (13)$$

其中, $P_r(d)$ 为接收功率; d 为传播距离; γ 为路径损耗指数。由此可以得到相对距离的估计公式:

$$d_{i,t} = \left(\frac{P_{\text{ref}}}{X_{i,t} + \rho} \right)^{\frac{1}{\gamma}} \quad (14)$$

其中, P_{ref} 为参考功率常数; ρ 为防止除零的微小量。 P_{ref} 与 γ 在此处不需要精确的物理测量值,而是作为归一化常数使用。其目的在于建立一种单调的映射关系,使得较大的接收强度对应较小的相对距离,反之亦然。最终得到的 $d_{i,t}$ 并不表示真实的物理距离,而是一种相对尺度上的距离度量。随后,距离偏移编码可以利用两个样本的相对距离差来计算:

$$\Delta d = d_1 - d_2 \quad (15)$$

3.2.2 速度偏移编码

在速度变化的编码方面,考虑到速度是距离随时间的变化率,对于样本 k 的距离向量 \mathbf{d}_k ,其速度可以估计:

$$\mathbf{v}_k = \frac{\partial \mathbf{d}_k}{\partial t} \quad (16)$$

在计算速度编码时,可采用有限差分来近似上式:

$$\mathbf{v}_{k,t} = \frac{\mathbf{d}_{k,t+1} - \mathbf{d}_{k,t}}{\Delta t} \quad (17)$$

其中, Δt 为相邻采样点的时间间隔。进一步地,两个样本之间的速度差异可通过下式计算:

$$\Delta \mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2 \quad (18)$$

最终,将距离差分 Δd 与速度差分 $\Delta \mathbf{v}$ 进行堆叠,得到完整的条件编码:

$$\mathbf{C} = \begin{bmatrix} \Delta \mathbf{v} \\ \Delta d \end{bmatrix} \in \mathbb{R}^{2 \times 49 \times N_{\text{time}}} \quad (19)$$

该编码既符合射频信号传播的物理规律,又刻画了样本之间在距离和速度上的相对变化,使得生成模型能够学习到更贴近实际的变化模式,为后续的识别模型训练奠定基础。

3.3 行为多样的数据合成

为了在有限的RFID感知数据基础上提高训练数据的多样性与分布广度,本文采用扩散模型进行数据合成。与传统扩散模型从纯噪声逐步生成目标样本的方式^[14]不同,本文的扩散过程以一个真实样本 X 为起点,并结合行为偏移编码引导生成过程。具体而言,首先从数据集中随机抽取两个属于同一类别的样本,通过计算其在距离与速度上的偏移来形成条件编码 \mathbf{C} ,然后利用该条件将样本 X 迁移至新的分布,从而生成具有动态变化特征的合成样本 X_{gen} 。这一策略不仅保证生成样本符合物理规律,同时也通过引入条件扰动提升了合成数据的多样性。

如图5所示,本文提出的生成模型的核心为轻量化的U-Net^[37]。U-Net由对称的下采样与上采样模块组成,其中下采样部分提取多尺度特征,上采样部分逐步恢复数据的时空分辨率。中间层通过引入时间步以及条件编码实现动态特征调制。时间嵌入部用正弦函数位置编码,其形式为

$$\text{PE}(t) = \begin{bmatrix} \sin(t \cdot \omega_1), \cos(t \cdot \omega_1), \sin(t \cdot \omega_2), \\ \cos(t \cdot \omega_2), \dots, \sin(t \cdot \omega_{\text{dim}}), \cos(t \cdot \omega_{\text{dim}}) \end{bmatrix} \quad (20)$$

其中, ω_k 为预设频率分量; dim 为嵌入维度。该编码提供了时间信息,使得模型能够学习不同扩散步数下的去噪特征。类别信息通过嵌入矩阵映射为向量,并与时间嵌入相加,经过多层感知机投影后注入至U-Net中间层,实现类别引导的生成控制。条件编码 \mathbf{C} 在输入阶段与样本特征拼接后送入网络。模型的整体输入为 $X_{\text{in}} = \text{Concat}(X_t, \mathbf{C})$,其中, X_t 为扩散过程在第 t 步的样本。

该模型的训练遵循正向加噪与反向去噪的过程。

在正向过程中,真实样本 X_0 逐步加入高斯噪声,得到第 t 步的样本:

$$X_t = \sqrt{\bar{\alpha}_t} X_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon \quad (21)$$

其中, $\epsilon \sim \mathcal{N}(0, \mathbf{I})$, $\bar{\alpha}_t = \prod_{s=1}^t (1 - \beta_s)$, β_s 为噪声调度参数。

最终,在第 T 步得到转换后样本 X_T 。

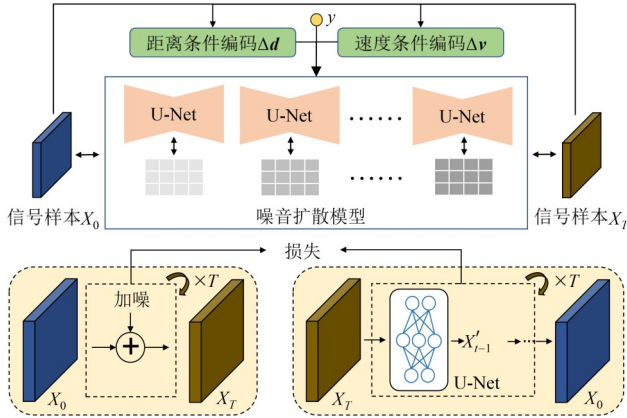


图5 行为多样性指导的数据合成模型

Figure 5 Signal synthesis model conditioned on behavioral diversity

在反向过程中,模型学习预测噪声 $\hat{\epsilon}_\theta(X_t, t, y, \mathbf{C})$,并通过最小化均方误差进行训练:

$$L(\theta) = \mathbb{E}_{X_0, \epsilon, t} \epsilon - \hat{\epsilon}_\theta(X_t, t, y, \mathbf{C})^2 \quad (22)$$

其中, y 表示类别标签。该损失函数保证模型能够在给定条件约束下重构噪声分布,从而实现受控生成。

在推理阶段,采样过程从观测样本 X 出发,通过条件扩散逐步迭代去噪,最终得到生成样本 X_{gen} :

$$X_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(X_t - \frac{\beta_t}{\sqrt{1 - \bar{\alpha}_t}} \hat{\epsilon}_\theta(X_t, t, y, \mathbf{C}) \right) + \sigma_t z \quad (23)$$

其中, $z \sim \mathcal{N}(0, \mathbf{I})$; σ_t 为噪声项。最终生成的样本在保持类别一致性的同时,体现了由条件编码引导的分布扩展,从而实现基于扩散模型的数据增强。

4 实验与结果

4.1 实验设置与度量

为验证所提出的数据增强方法的有效性,本章在真实物理环境中进行了实验。如图6所示,实验采用49个Alien-9629标签构建标签阵列,用于手势数据采集;使用配备Laird A9028天线的Impinj R420读写器完成射频信号的发射与接收。数据预处理由搭载Intel(R) Core(TM) i5-8400 CPU的Dell主机完成,数据生成模型的训练与测试则在配备8 GB显存的NVIDIA GeForce GTX 1070 GPU上执行。模型参数设置如下:默认情况下,式(9)中窗口长度 W 为3,检测

阈值 θ 为3;式(13)中的参考功率 P_{ref} 为损失系数, γ 为2,微小量 ρ 为 1×10^{-6} 。生成模型的学习率为 1×10^{-4} 。

为进行数据采集,实验共邀请了5名年龄范围在21~31岁之间的志愿者(2名女性和3名男性)在标签阵列前进行手势录入。选取常用的10个阿拉伯数字0到9作为目标手势。每个受试者对每个手势至少录入50次,共收集超过4 000个信号样本。每个样本的数据维度为 49×30 ,即每个标签选取连续的30个采样点以覆盖整个手势录入过程,其中每个元素表示该标签在对应时刻获得的信号强度值。

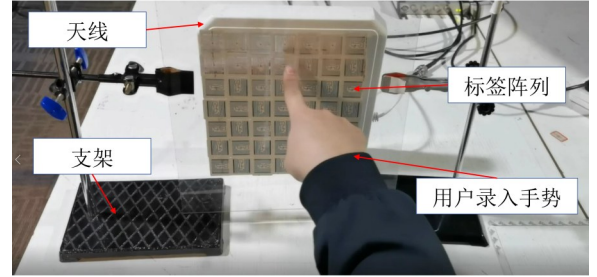


图6 物理世界的实验设置

Figure 6 Real-world experiment setup

本章选取了两个定量指标来全面评估所提出数据增强方法的性能,即结构相似性指标 (Structural Similarity Index Measure, SSIM)^[38]和准确率 (Accuracy)^[39]。其中,SSIM是一种用于衡量两组数据在亮度、对比度以及结构特征方面相似程度的指标。该指标通过分析样本间的均值、方差及协方差等统计特征来计算其相似性,取值范围为 $[0, 1]$,数值越接近1表示两者越相似。本章利用SSIM度量生成样本与真实样本之间的结构一致性,从而评估生成数据的保真度。准确率则用于衡量识别模型在预测未知样本类别时的准确程度,反映了增强数据对模型判别能力提升的有效性。其取值范围同样为 $[0, 1]$,数值越高表示模型的分性能越优。通过综合分析SSIM与准确率,可以从数据层面和任务层面对所提出方法的真实性与实用性进行双重验证。

4.2 总体性能

为了全面评估所设计的条件指导式生成模型在数据合成上的保真度与区分性,本节首先采用SSIM对生成数据与真实数据进行量化对比。SSIM能够从结构信息、亮度和对比度等多个维度刻画两组数据的相似程度,因此在衡量生成数据质量时具有较强的解释性和可信度。该实验不仅计算了类内样本(即生成样本与同一类别真实样本之间)SSIM,还计算了类间样本(即生成样本与不同类别真实样本之间)的SSIM,从而同时考察数据的保真度与可区分性。实

验结果如图7所示,可以看到生成数据与真实数据之间整体的SSIM值非常高,表明模型具备良好的拟合能力。其中,类内样本的平均SSIM高达0.857 2,说明生成的数据在结构层面与真实数据十分相似,具备极高的保真度^[40]。相比之下,类间样本的平均SSIM为0.844 8,数值略低于类内样本。这一结果反映了两方面意义:一方面,生成数据与真实数据整体上高度一致,充分保证了信号的真实性和真实性;另一方面,类间相似度的下降说明不同类别之间仍保持明显差异,从而保证了数据在分类和感知任务中具有较好的可分性。进一步地,SSIM值的整体偏高也体现了条件指导策略在生成过程中有效约束了信号特征,使得合成样本既真实又具有辨识度。综上,实验验证了本方法在数据生成的“真实性—区分性”平衡上表现优异,后续基于合成数据增强的手势识别和泛化研究提供了可靠保障。

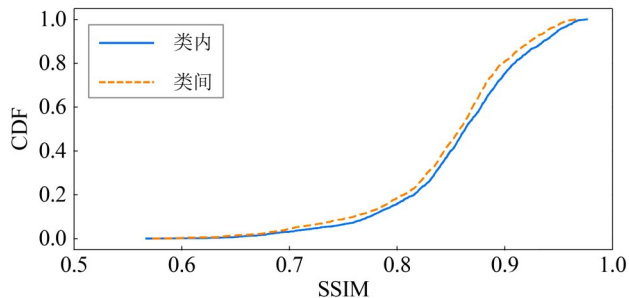


图7 真实样本与生成样本之间的SSIM分布

Figure 7 SSIM distributions between real and synthesized samples

接着,本节进一步评估了在利用合成数据进行增强之后,所带来的识别准确率提升。实验的核心目标在于验证本文提出的条件指导式数据生成方法在增强有限数据集时的有效性和优越性。为了进行公平对比,本节复现了两种最新的针对RFID感知的数据增强方法: FewSense^[12]和RF-ACCLDM^[14]。其中, FewSense是一类基于传统数据变换的增强方法,通过旋转、倒置等手段生成新的样本,具有实现简单的优点,但其生成数据的多样性有限。与之相比,RF-ACCLDM是一种基于条件扩散模型的生成方法,它在类别条件的约束下进行数据合成,能够一定程度上增加样本的多样性。为了体现增强方法在不同类型分类器上的普适性,本实验分别在LR^[34]这一经典的浅层传统分类器,以及CNN^[41]这一代表性的深度学习模型上进行测试。实验设置如下:在“未增强”的基线情况下,仅使用10%的真实数据作为训练集,其余90%的真实数据作为测试集;在“增强”的情况下,训练集由10%的真实数据和合成数据共同组成,测试集仍为未参与识别模型训练的真实数据。

图8给出了不同方法在这两类分类器上的识别

准确率对比结果。该结果表明,当不进行增强时,LR和CNN的识别准确率分别为61.7%和71.0%,处于较低水平,表明有限的训练样本不足以支撑鲁棒的识别性能。当采用本文提出的方法进行数据增强后,LR和CNN的准确率分别大幅提升至88.1%和93.1%,增益均超过20个百分点。这表明本文方法不仅显著扩展了训练数据的数量,更重要的是有效提升了数据分布的行为多样性,使得模型能够更好地捕捉潜在特征,从而具备更强的泛化能力。相比之下, FewSense增强后的准确率为40.0%(LR)和78.0%(CNN),在CNN上表现出一定提升,但在LR上则出现了显著下降,说明其增强方式并不稳健。RF-ACCLDM增强后的准确率为62.5%(LR)和52.7%(CNN),结果同样呈现出两极化趋势:在LR上略有提升,而在CNN上却出现了性能的退化。这一现象说明,现有方法在扩展数据规模的同时,并未充分保证数据分布的合理性,甚至可能生成了部分与物理规律不符的样本,从而对模型训练造成干扰。综上所述,本文提出的RFID数据增强方法不仅能够有效提升生成数据的保真度,还能够在引入物理先验(距离条件和速度条件)的指导下显著增加数据分布的多样性,从而在不同类型的分类器上均取得了稳定而显著的性能提升。这验证了本文方法在有限训练数据条件下的实用价值,并显示出其在RFID感知任务中相较于现有增强方法的明显优势。用户可以利用虚拟数据代替真实数据作为训练集,有效降低数据采集成本。

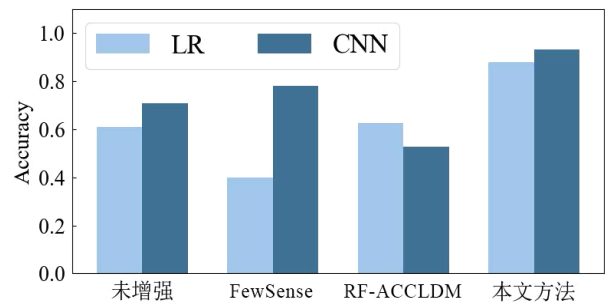


图8 总体数据增强效果

Figure 8 Overall effectiveness of data augmentation

4.3 消融实验

为了进一步验证所提出的面向行为多样性的数据生成方法在扩展数据分布和提升模型泛化能力方面的有效性,本实验系统地对比了不同行为条件指导下的识别性能表现。具体而言,分别仅使用距离偏移编码与标签,或仅使用速度偏移编码与标签作为生成条件,对生成样本进行指导,并评估其在识别任务中的准确率表现。该设计旨在探究单一行为条件对生成样本分布与模型性能的具体贡献。实验结果如图9

所示。结果表明,在仅使用单一行为条件约束的情况下,无论采用 LR 模型还是 CNN 模型,识别准确率均较未使用行为条件时有显著提升,表明引入行为多样性能够有效改善生成样本的质量。进一步比较发现,距离条件与速度条件带来的性能增益相对接近,说明二者在刻画信号传播特性和动作差异方面具有互补作用。当同时引入距离与速度两种偏移编码时,识别准确率再次显著提高,验证了多行为条件联合指导在提升生成样本多样性和表达能力方面的协同效应。综上所述,本文提出的行为条件指导机制不仅能够生成阶段有效引导模型合成更加丰富、分布更广泛的样本,还能在下游识别任务中显著提升模型对未知样本的泛化能力和稳健性,充分体现了行为多样性人机交互感知应用中的重要作用。

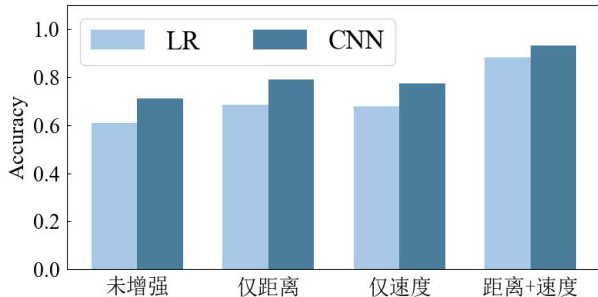


图9 不同行为偏移约束下的识别性能

Figure 9 Recognition performance conditioned on different behavioral shift

4.4 合成数据的比例影响

为了系统分析合成数据与真实数据比例对识别性能的影响,本实验以合成数据规模为变量,采用5倍为步长,将合成数据相对于真实数据的倍数从0逐步增加至70,并分别在 LR 与 CNN 模型上评估识别准确率的变化情况。该实验旨在探讨数据生成在不同规模下对模型性能的促进作用,从而确定生成数据在增强识别能力方面的最优使用范围。实验结果如图 10 所示。可以观察到,随着合成数据比例的增加,模型的识别准确率持续上升,表明生成数据在丰富样本多样性和提升模型泛化能力方面具有显著作用。从曲线的变化趋势来看,准确率的提升呈现“先快后慢”的特征:在比例较低阶段(0~30倍)时,准确率迅速增长;而当比例继续增大后,增长幅度逐渐减缓,最终趋于饱和。这一现象说明,当生成数据的数量超过一定阈值后,其对模型性能的边际贡献逐渐降低。综合分析可知,适量的合成数据即可显著增强模型的识别性能,而过量生成可能导致样本的冗余与分布重叠。因此,在实际应用中,无需依赖大规模的合成数据,即可实现性能与计算成本之间的平衡,从而有效提升无线感知系统的整体效率与可扩展性。

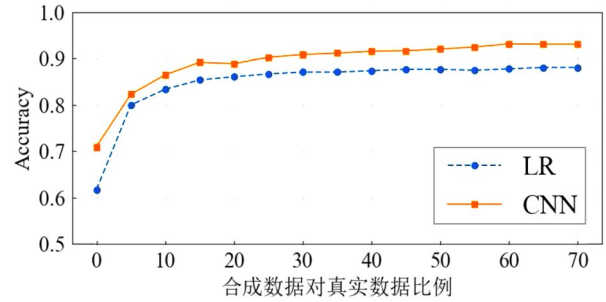


图10 不同合成数据对真实数据比例下的准确率

Figure 10 Accuracy under different ratios of synthetic data to real data

4.5 感知任务规模的影响

为了深入分析待识别类别规模对数据增强性能的影响,本实验将手势类别的数量以步长1从5逐步增加至10,并在不同类别规模下分别评估了 LR 与 CNN 分类器的识别准确率。实验结果如图 11 所示,展示了数据增强前(Before)与数据增强后(After)两种情况下的准确率变化趋势。可以观察到,随着手势类别数量的增加,无论在增强前还是增强后,LR 和 CNN 模型的识别准确率均呈现出逐渐下降的趋势。这一现象是合理的,因为更大的类别集合会引入更复杂的特征分布与更高的区分难度,从而增加模型的分类挑战。然而,值得注意的是,尽管总体准确率有所下降,数据增强带来的性能提升仍然显著,说明所提出的物理条件指导式增强方法能够在更复杂的分类任务中依然保持较强的有效性。

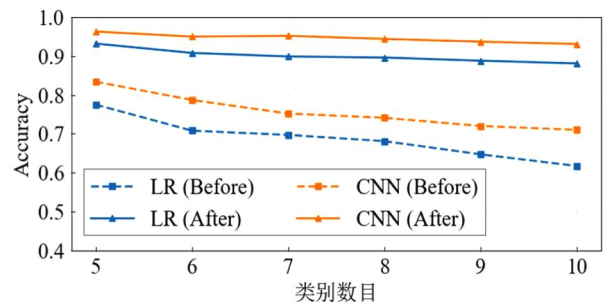


图11 类别规模变化的影响

Figure 11 Impact of class scale

为进一步量化这种性能增益,表1列出了在不同类别规模下的增强前后准确率差值,即数据增强带来的增益。

从结果可以看到,随着类别数量的增加,增益值呈现持续上升的趋势,这表明在类别数量增多、样本分布更复杂的情况下,数据增强所带来的效果更加显著。综上,本文提出的增强方法不仅能在小规模分类任务中提升识别性能,更有望在大类别、高复杂度的识别任务中展现出强适应性与潜在优势。

表1 数据增强在LR和CNN上的增益

Table 1 Performance gains of data augmentation on LR and CNN

类别数目	LR/%	CNN/%
5	15.7	12.9
6	20.0	16.3
7	20.2	20.0
8	21.5	20.3
9	24.1	21.7
10	26.4	22.1

4.6 应用扩展验证

为验证所提出的数据增强方法在行为识别^[42]之外其他RFID感知任务中的可扩展性,本实验进一步开展了身份识别性能评估。由于数据采集阶段共记录了5名参与者的样本,因此可构建5类身份识别实验。实验设置与手势识别任务保持一致:在未增强条件下,随机选取10%的真实数据作为训练集,其余为测试集;在增强条件下,训练集由10%的真实数据与生成数据共同构成,测试集仍为剩余真实数据。实验结果如图12所示。可以看出,无论在LR还是CNN模型上,数据增强均显著提升了身份识别的准确率,表明本文方法在其他RFID感知任务中具有良好的通用性与推广潜力。

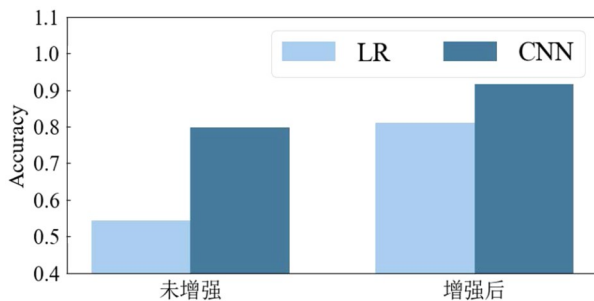


图12 身份识别应用的数据增强效果

Figure 12 Effectiveness of data augmentation on user identification

4.7 开销评估

为评估本文所提出数据生成方法在实际应用中的存储与时间开销,本实验在配备8 GB显存的NVIDIA GeForce GTX 1070 GPU平台上,对生成模型的数量规模以及单条样本的生成耗时进行了统计。实验结果表明,该生成模型的数量约为318 M,合成一条虚拟样本的平均耗时约为1.2 s;同时,将生成模型保存至本地约占用12 MB的存储空间。上述结果表明,本文的方法在保证生成质量的同时,对计算与存储资源的需求相对可控,能够在中等配置的硬件条件下稳定运行,具有较好的工程可部署性与实用性,为在真实场景中进行大规模数据增强提供了可行支撑。此外,由于生成过程可离线完成,合成数据可

直接用于后续识别模型训练,因此不会对实际感知系统的在线运行带来额外负担。

5 结论

本文针对RFID手势识别中训练数据分布受限、模型泛化能力不足的问题,提出了一种面向行为多样性的数据合成方法,实现数据增强并提升识别模型的泛化性能。具体而言,首先通过对RFID信号传播过程的建模与分析,验证了人手与标签之间的距离及运动速度是决定手势样本差异的关键物理因素;随后,将距离与速度量化为条件编码并嵌入扩散模型,使生成过程能够在不同物理约束下实现分布可控的样本扩展;最后,通过行为条件引导的高保真数据合成,构建出兼具多样性且符合实际传播规律的训练数据集,从而显著增强手势识别模型在未知样本上的识别能力。通过合成数据替代部分真实数据,本文的方法可有效降低用户的数据采集开销。实验结果表明,该方法在真实场景中识别准确率较基线模型提升超过20%,显著优于传统数据增强策略。同时,该方法可有效扩展至其他人机交互应用中。

本文通过引入距离与速度偏移量来提升生成数据的行为多样性。然而,在实际应用中,人手的其他行为因素(如旋转角度等)同样可能导致同类样本之间的显著差异。未来的研究可进一步引入自适应的行为条件编码机制,使模型能够根据输入信号特征自动、全面地学习行为偏移规律。例如,可结合深度学习与注意力机制,从历史样本中自动挖掘对信号分布影响显著的关键行为变量,并将其嵌入生成模型的条件空间中,从而实现对多维行为变化的自适应数据增强。

参考文献

- [1] Xiao Rui, Liu Jianwei, Han Jinsong, et al. OneFi: One-shot recognition for unseen gesture via COTS WiFi[C]//Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems. New York: ACM, 2021: 206-219.
- [2] 谷军霞, 丁晓青, 王生进. 基于人体行为3D模型的2D行为识别[J]. 自动化学报, 2010, 36(1): 46-53.
Gu Junxia, Ding Xiaoqing, Wang Shengjin. Human 3D model-based 2D action recognition[J]. Acta Automatica Sinica, 2010, 36(1): 46-53. (in Chinese)
- [3] 曹媛媛, 黄飞跃, 陶霖密, 等. 自然环境下日常动作的在线识别[J]. 电子学报, 2009, 37(S1): 16-21.
Cao Yuanyuan, Huang Feiyue, Tao Linmi, et al. On-line recognition of actions in daily living[J]. Acta Electronica Sinica, 2009, 37(S1): 16-21. (in Chinese)
- [4] Zou Yongpan, Xiao Jiang, Han Jinsong, et al. GRfid: A device-free RFID-based gesture recognition system[J]. IEEE Transactions on Mobile Computing, 2017, 16(2): 381-393.

- [5] Liu Jianwei, Zou Xiang, Lin Feng, et al. Hand-key: Leveraging multiple hand biometrics for attack-resilient user authentication using COTS RFID[C]//2021 IEEE 41st International Conference on Distributed Computing Systems. Piscataway: IEEE, 2021: 1042-1052.
- [6] Liu Jianwei, Cui Kaiyan, Zou Xiang, et al. Reliable multi-factor user authentication with one single finger swipe[J]. IEEE/ACM Transactions on Networking, 2023, 31(3): 1117-1131.
- [7] Zhang Shigeng, Ma Zijing, Yang Chengwei, et al. Real-time and accurate gesture recognition with commercial RFID devices[J]. IEEE Transactions on Mobile Computing, 2023, 22(12): 7327-7342.
- [8] Wang Chuyu, Liu Jian, Chen Yingying, et al. Multi - touch in the air: Device-free finger tracking and gesture recognition *via* COTS RFID[C]//IEEE INFOCOM 2018-IEEE Conference on Computer Communications. Piscataway: IEEE, 2018: 1691-1699.
- [9] Gong Chen, Liang Bo, Gao Wei, et al. Data can speak for itself: Quality-guided utilization of wireless synthetic data[C]//Proceedings of the 23rd Annual International Conference on Mobile Systems, Applications and Services. New York: ACM, 2025: 209-222.
- [10] Zhao Leqi, Xiao Rui, Liu Jianwei, et al. One is enough: Enabling one-shot device-free gesture recognition with COTS WiFi[C]//IEEE INFOCOM 2024-IEEE Conference on Computer Communications. Piscataway: IEEE, 2024: 1231-1240.
- [11] Jiang Wei, Zhang Kai, Wang Nan, et al. MeshCut data augmentation for deep learning in computer vision[J]. PLoS One, 2020, 15(12): e0243613.
- [12] Zhang Xiaobin, Xu Hongzhe, Gao Wei, et al. FewSense: Enabling few-shot gesture recognition *via* COTS RFID[J]. ACM Transactions on Sensor Networks, 2025, 21(5): 1-14.
- [13] Yang Chao, Wang Ziqi, Mao Shiwen. RFPose-GAN: Data augmentation for RFID based 3D human pose tracking[C]//2022 IEEE 12th International Conference on RFID Technology and Applications. Piscataway: IEEE, 2022: 138-141.
- [14] Wang Ziqi, Mao Shiwen. AIGC for wireless sensing: Diffusion-empowered human activity sensing[J]. IEEE Transactions on Cognitive Communications and Networking, 2025, 11(2): 657-671.
- [15] Savin J, Gaudez C, Gilles M A, et al. Evidence of movement variability patterns during a repetitive pointing task until exhaustion[J]. Applied Ergonomics, 2021, 96: 103464.
- [16] Li Tianxing, Liu Qiang, Zhou Xia. Practical human sensing in the light[C]//Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services. New York: ACM, 2016: 71-84.
- [17] Wang Minsi, Ni Bingbing, Yang Xiaokang. Recurrent modeling of interaction context for collective activity recognition[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 7408-7416.
- [18] Withana A, Peiris R, Samarasekara N, et al. zSense: Enabling shallow depth gesture recognition for greater input expressivity on smart wearables[C]//Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. New York: ACM, 2015: 3661-3670.
- [19] Zhao Tianming, Liu Jian, Wang Yan, et al. PPG-based finger-level gesture recognition leveraging wearables[C]//IEEE INFOCOM 2018 - IEEE Conference on Computer Communications. Piscataway: IEEE, 2018: 1457-1465.
- [20] Kalgaonkar K, Raj B. One-handed gesture recognition using ultrasonic Doppler sonar[C]//2009 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE, 2009: 1889-1892.
- [21] 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, 2012: 341-350.
- [22] Cao Dian, Wang Dong, Zhang Qian, et al. Towards domain-independent complex and fine-grained gesture recognition with RFID[J]. Proceedings of the ACM on Human-Computer Interaction, 2020, 4: 1-22.
- [23] Ma Zijing, Zhang Shigeng, Liu Jia, et al. RF-Siamese: Approaching accurate RFID gesture recognition with one sample[J]. IEEE Transactions on Mobile Computing, 2024, 23(1): 797-811.
- [24] Hou Weiyang, Wu Chenshu. RFBoost: Understanding and boosting deep WiFi sensing *via* physical data augmentation[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2024, 8(2): 1-26.
- [25] Mikołajczyk A, Grochowski M. Data augmentation for improving deep learning in image classification problem[C]//2018 International Interdisciplinary PhD Workshop. Piscataway: IEEE, 2018: 117-122.
- [26] Oneață D, Cucu H. Improving multimodal speech recognition by data augmentation and speech representations[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2022: 4578-4587.
- [27] Yang Ling, Zhang Zhilong, Song Yang, et al. Diffusion models: A comprehensive survey of methods and applications[J]. ACM Computing Surveys, 2024, 56(4): 1-39.
- [28] He Yuan, Sun Yimiao, Guo Xiuzhen. RF computing: A new realm of IoT research[J]. Journal of Computer Sci-

- ence and Technology, 2025, 40(4): 941-956.
- [29] 陆琪鹏, 刘亚丽, 刘长庚, 等. 基于区块链的RFID供应链产品所有权转移方案[J]. 电子学报, 2025, 53(2): 451-459. Lu Qipeng, Liu Yali, Liu Changgeng, et al. Product ownership transfer scheme of RFID-enabled supply chain based on blockchain[J]. Acta Electronica Sinica, 2025, 53(2): 451-459. (in Chinese)
- [30] Finkenzerler K. RFID Handbook: Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and near-Field Communication[M]. Hoboken: Wiley, 2010.
- [31] Xu Huanyuan, Zhang Yajun, Yang Zhixiong, et al. RFCSign: A Chinese sign language recognition system based on large kernel convolution and normalization-based attention[J]. IEEE Access, 2023, 11: 133767-133780.
- [32] Rappaport T S. Wireless Communications--Principles and Practice, (The Book End)[J]. Microwave Journal, 2002, 45(12): 128-129.
- [33] McInnes L, Healy J, Melville J. UMAP: Uniform manifold approximation and projection for dimension reduction[PP/OL]. V3.arXiv (2020-09-18)[2026-01-02]. <https://doi.org/10.48550/arXiv.1802.03426>.
- [34] Ng A Y, Jordan M I. On discriminative vs. generative classifiers: A comparison of logistic regression and naive Bayes[C]//Proceedings of the 15th International Conference on Neural Information Processing Systems: Natural and Synthetic. New York: ACM, 2001: 841-848.
- [35] Arachchige C N P G, Prendergast L A. Confidence intervals for median absolute deviations[J]. Communications in Statistics-Simulation and Computation, 2026, 55(1): 13-22.
- [36] Henderi H. Comparison of Min-Max normalization and Z-score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer[J]. IJIS: International Journal of Informatics and Information Systems, 2021, 4(1): 13-20.
- [37] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[M]//Medical Image Computing and Computer-Assisted Intervention MICCAI 2015. ChamSpringer International Publishing, 2015: 234-241.
- [38] Wang Zhou, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [39] Liu Jianwei, Yuan Jiantao, Yu Guanding, et al. Efficient one-shot gesture recognition for WiFi ISAC via aug-meta learning[J]. IEEE Journal on Selected Areas in Communications, 2025, 43(11): 3766-3781.
- [40] Chi Guoxuan, Yang Zheng, Wu Chenshu, et al. RF-diffusion: Radio signal generation via time-frequency diffusion[C]//Proceedings of the 30th Annual International Conference on Mobile Computing and Networking. New York: ACM, 2024: 77-92.
- [41] Li Zewen, Liu Fan, Yang Wenjie, et al. A survey of convolutional neural networks: Analysis, applications, and prospects[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(12): 6999-7019.
- [42] 卿宇寒, 高陈强, 谭卓林, 等. 基于时空自适应融合的双模行为识别[J]. 电子学报, 2025, 53(7): 2389-2400. Yuhan Qing, Gao Chenqiang, Tan Zhuolin, et al. Bimodal action recognition based on spatiotemporal adaptive fusion[J]. Acta Electronica Sinica, 2025, 53(7): 2389-2400. (in Chinese)



刘建伟 男,1997年1月出生于甘肃省定西市。现为浙江大学博士后。在国内外发表学术论文40余篇。研究方向包括无线感知、移动计算、深度学习等。

E-mail: jianweiliu@zju.edu.cn



姚信威 男,1986年5月出生于浙江省诸暨市。现为浙江工业大学先进技术研究院副院长,计算机科学与技术学院教授、博士生导师。研究方向包括群智计算、群智感知与协同、智能物联网等。

E-mail: xwyao@zjut.edu.cn



陈佳彤 男,2002年5月出生于重庆市。现为浙江大学硕士生。研究方向包括无线感知和深度学习。

E-mail: chenjiatong@zju.edu.cn



韩劲松 男,1975年12月出生于山东省淄博市。现为浙江大学计算机科学与技术学院教授、博士生导师。研究方向包括物联网安全、智能感知、人工智能安全、隐私保护等。

E-mail: hanjinsong@zju.edu.cn