**文章编号：**

基于双通道特征融合的Wi-Fi CSI人体行为识别

卫嘉琪１，陈亮1，梁颖1

(1.西安工程大学计算机科学学院 陕西 西安 710048；

**摘要：**随着智慧医疗与智能家居的快速发展，基于Wi-Fi信道状态信息 (Channel State Information, CSI) 的无设备人体行为识别 (Human Activity Recognition, HAR) 技术被广泛研究与应用，有效解决了传统传感器部署繁琐与隐私泄露的问题，但其识别性能仍易受环境噪声与多径干扰影响，导致识别稳定性不足、细粒度行为难以区分。为此，本文提出一种融合相位与振幅特征的双通道Transformer模型PA-TCNet (Phase-Amplitude Time-Channel Network)。该模型采用双通道架构分别处理振幅与相位信息，并引入可学习相对位置编码 (Learnable-Relative Positional Embedding, LPE) 与相对位置注意力机制，以增强模型对时序与空间特征的鲁棒表征能力。为进一步挖掘两类特征的深层耦合关系，我们引入Kolmogorov–Arnold Network (KAN) 实现非线性特征融合，从而提升对相似动作的区分能力。实验结果表明，PA-TCNet在多个公开数据集上均取得优异性能，显著降低了环境干扰带来的识别误差，为复杂室内场景下的高精度行为识别提供了可靠解决方案，在智慧医疗监测、智能家居交互等领域具有广泛的应用潜力。

**关键词：**Wi-Fi感知；信道状态信息；行为识别；相位-幅度特征；双通道网络；相对位置注意力；

**中图分类号：**TP3　　　　**文献标志码：**A　　 **DOI:**

RESEARCH ON A GRAPH RETRIEVAL-AUGMENTED GENERATION METHOD FOR QUESTION ANSWERING IN THE POWER GRID DOMAIN

Wei Jiaqi 1, Chen Liang 1, Liang Ying2

(1.School of Computer Science, Xi’an Polytechnic University, Xi’an 710048 ,China；

2.School of Computer Science, Xi’an Polytechnic University, Xi’an 710048 ,China)

**Abstract:** To improve the accuracy and robustness of Wi-Fi CSI-based human activity recognition (HAR), this paper proposes a dual-channel transformer model named PA-TCNet (Phase-Amplitude Time-Channel Network), which fuses amplitude and phase features. Combining learnable relative positional embedding (LPE) and the relative position attention mechanism, PA-TCNet effectively mitigates the impacts of environmental noise and multipath interference on traditional Wi-Fi HAR methods. By introducing the Kolmogorov–Arnold Network (KAN) nonlinear fusion mechanism, the model captures the complex coupling relationship between amplitude and phase, thus enhancing its ability to distinguish fine-grained activity differences. Experimental results show that PA-TCNet achieves excellent performance on multiple datasets, especially in differentiating similar activities. The model design improves HAR accuracy in complex scenarios and has broad application potential, particularly in smart healthcare and smart home domains.

**Keywords**: Wi-Fi , Channel State Information, Activity Recognition, Phase and Amplitude, attention mechanism

# 引 言

人体行为识别(Human Activity Recognition, HAR) 是一项通过技术手段自动识别与解析人类活动的关键技术。在智慧医疗、智能家居及养老看护等场景快速发展的背景下，基于Wi-Fi信道状态信息(CSI)的无设备感知技术应运而生。该技术利用无处不在的Wi-Fi信号，通过分析人体活动对无线信道造成的细微扰动来实现感知。相较于依赖摄像头或可穿戴设备的方法，该技术具有非侵入性、成本低廉且天然保护用户视觉隐私的显著优势，为实现连续、自然、无感的行为监测提供了理想方案。通过对日常活动（如行走、坐下）与异常状态（如跌倒）的精准识别，该技术能为远程健康监护、个性化生活辅助以及智能环境交互提供核心数据支持，对于提升生活安全性、便利性与照护效率具有重要的现实意义，展现出广阔的应用前景。

现阶段，基于Wi-Fi信道状态信息的行为识别方法主要沿两个方向发展，首先在特征利用方面，大多数研究仍以振幅特征为核心，Jaeseong Son等人提出对CSI振幅数据进行专门的预处理与特征增强，通过非线性变换将其扩展为三通道向量以放大特征差异。然而，这种方法对蕴含丰富路径动态信息的相位特征利用仍然不足，并且缺乏能够有效刻画振幅与相位间复杂、非线性耦合关系的融合机制，从而限制了模型对细微动作差异的感知能力。其次，在时空建模方面，为更有效地捕捉CSI数据中的长程依赖，Transformer架构及其变体被引入该领域。M. D. I. Kobir等人提出了一种混合深度学习方法，通过整合信号预处理、定向数据增强以及定制化的CNN与Transformer模型来提升性能。然而，这类方法在强化时序建模能力的同时，通常将CSI的子载波维度（通道维度）与时间维度进行扁平化或独立处理，未能显式地建模时间与通道双维度之间复杂的交互与协同依赖关系。正是这种多维关联的缺失，制约了模型对“坐下”与“站起”等高度相似行为的精细区分能力。

为解决上述问题，我们提出了一种名为PA-TCNet (Phase-Amplitude Time-Channel Network) 的双通道Transformer模型。该模型通过融合相位与振幅的双流特征、引入可学习相对位置编码与注意力机制，以及应用KAN进行非线性融合，旨在提升对细粒度相似行为的识别精度与鲁棒性。首先，我们在特征层面引入相位信息，与振幅特征共同构成双流输入，以充分利用相位所蕴含的丰富路径动态，为识别提供更全面的时频域特征基础。其次，我们设计了双流Transformer架构，通过可学习相对位置编码（LPE）和通道维度的相对位置注意力（MRPT），显式地协同建模时间演变与多子载波（通道）间的复杂依赖关系，从而增强对“坐下”与“站起”等相似动作的细微差异的捕捉能力。最后，我们引入了基于Kolmogorov–Arnold Network (KAN) 的融合机制（KAN-GRN）。该机制通过可学习的基函数，自适应地刻画振幅与相位特征间的高阶非线性耦合，实现更精细、更稳健的特征融合。我们在StanWiFi、MultiEnv和自建SimAct数据集上进行了验证。实验结果表明，我们的模型达到了优异的效果：在StanWiFi数据集上识别率达到99.83%；在MultiEnv的办公室和大厅两种差异显著的场景下，识别率分别达到99.13% 和98.67%，展现了强大的环境鲁棒性

综上所述，本文主要贡献如下:1.引入相位信息，增强特征表达，在传统振幅特征表征的基础上，引入相位信息作为补充特征，并通过统一的特征提取策略实现振幅与相位特征的深度融合，不仅丰富了CSI信号的表征维度，还通过相位信息提供的额外时频域特征，显著提升了模型对细微行为差异的识别能力，为Wi-Fi相似行为识别提供了更加全面的特征基础。

2.提出双流Transformer架构，增强长距离依赖建模，通过引入可学习相对位置编码LPE和相对位置注意力机制，有效捕捉CSI信号在时间和通道维度上的长距离依赖关系，提升了模型对复杂行为模式的表征能力。3.引入KAN进行非线性特征融合，用于刻画幅度与相位特征间的非线性耦合关系。该机制通过可学习的基函数展开，使网络能够自适应地逼近高阶映射，从而捕捉复杂的动态特征依赖。

# 相关工作

## 基于WIFI CSI行为识别发展

基于Wi-Fi信道状态信息（CSI）的无设备人体行为识别（HAR）技术，因其非侵入、低成本、保护隐私及穿透性等优势，已成为普适感知领域的研究热点。该技术通过解析人体活动对无线信道多径传播造成的细微扰动来实现识别，早期研究主要依赖于传统机器学习或浅层神经网络模型。随着深度学习的发展，卷积神经网络（CNN）和循环神经网络（RNN，尤其是LSTM）被广泛应用于该领域，显著提升了性能。CNN擅长提取局部空间（如子载波间）或时频特征，而LSTM则专注于建模时间序列依赖关系。例如，Wang等人[5]的工作表明，即使单独利用相位信息也能获得较高识别率，但其对噪声和环境变化的鲁棒性有限。Chen等人[6]则通过引入注意力机制的ABLSTM模型，进一步优化了时间特征的提取。为了同时捕捉CSI数据中的时空模式，研究者们进行了多种探索。一种主流思路是将CSI的时序和多子载波维度视为二维图像，进而采用2D-CNN进行处理[7]。这种方法虽能联合利用时空信息，但其卷积核的局部性限制了其对长程时间依赖的建模能力。近年来，Transformer架构凭借其全局注意力机制，在捕获长序列依赖关系方面展现出卓越性能，被成功引入Wi-Fi HAR任务中[9, 10]。例如，Yang等人的WiTransformer[10]提升了模型的鲁棒性与精度。为进一步增强模型适应性，Li等人[11]提出的THAT模型结合了卷积与Transformer，并利用高斯编码应对动作速度变化等问题。尽管上述方法取得了显著进展，但仍存在两个关键局限：其一，在特征层面，多数工作仍以振幅特征为绝对核心，对蕴含丰富路径动态信息的相位特征利用不足，且缺乏有效机制来深入挖掘振幅与相位之间复杂的非线性耦合关系；其二，在建模层面，现有方法往往将CSI的时间维与子载波（通道）维进行扁平化处理或独立建模，未能显式地刻画二者之间复杂的协同与交互依赖关系，这制约了模型对“坐下”、“站起”等细粒度相似动作的区分能力。本研究正是针对这两点不足，提出新的解决方案。

## KAN网络发展

在深度学习的架构演进中，多层感知机（MLP）凭借其通用近似能力成为绝大多数模型的基础组件。然而，MLP通常使用固定的非线性激活函数（如ReLU、Sigmoid），其参数主要存在于线性权重矩阵中，这种结构在逼近复杂的高维函数时可能效率不高，且可解释性较弱。近年来，一种新颖的网络架构—KAN网络(Kolmogorov–Arnold Network, KAN)被提出,其设计灵感源于柯尔莫哥洛夫-阿诺德表示定理，该定理指出任何多元连续函数都可以分解为有限个一元函数与加法运算的组合。这为使用一系列简单的、可学习的一元函数来构建复杂函数提供了理论基础。基于这一理论，KAN进行了一次根本性的架构创新，它将传统MLP中权重参数与固定激活函数的组合，替换为可学习的一元函数。具体而言，KAN层中的每条“连接”不再是一个简单的标量权重，而是一个可参数化的一元函数（如使用样条或其它基函数），网络的非线性主要由这些可学习的函数提供，而“加法”运算则替代了线性变换。这种核心设计的改变为KAN带来了一系列引人注目的潜在优势。首先，在表示能力方面，KAN理论上能以更小的网络规模、更高的精度逼近复杂函数，尤其擅长捕捉数据中光滑或具有特定结构的非线性关系。其次，在参数效率方面，相较于达到同等精度的MLP，KAN往往需要更少的参数，这得益于其结构更直接地匹配了函数的数学分解形式。最后，在可解释性方面，由于学习到的一元函数可以被可视化，研究者能够直观地分析不同输入变量如何通过特定的函数变换影响输出，这为理解模型内部工作机制打开了一扇新窗口。自提出以来，KAN已在多个科学计算和基础模型任务中展示了其潜力，例如，在复杂系统建模领域，Zhuoran Yang等人将KAN与图神经网络结合，提出了GGNKAN和GKAN模型，利用KAN强大的非线性拟合能力来同时重建网络连接与动力学，在相互作用恢复和预测误差上超越了现有方法。在更具挑战性的感知任务中，KAN也显示出独特价值，如Qiong Yang等人为应对水下复杂噪声干扰，构建了一个融合注意力机制与KAN的双通道识别模型，通过KAN刻画特征间的复杂非线性关系，显著提升了低信噪比环境下的目标识别准确率。此外，KAN的潜力也激发了硬件设计层面的探索，Yuhao Liu等人便着眼于其与传统人工神经网络不同的计算范式，开展了面向边缘设备的轻量化KAN硬件加速器前瞻性研究，以应对其部署挑战。这些研究共同表明，KAN不仅是一个有理论吸引力的模型，更在解决科学计算、鲁棒感知等实际问题上展现出切实有效的应用潜能。尽管将KAN应用于无线感知与特征融合领域的公开研究尚处起步阶段，但其强大的非线性拟合能力和结构特性，为解决传统融合方法（如简单拼接或线性加权）难以刻画振幅与相位间复杂、高阶耦合关系的瓶颈问题，提供了一个新的思路。因此本文将KAN引入Wi-Fi HAR的特征融合阶段，旨在构建一个能够自适应学习并表征幅相深度非线性交互的融合模块，以提升对细粒度动作的判别力。

## 无线感知技术的应用场景拓展

基于Wi-Fi CSI的无设备感知技术，其核心价值在于能够非侵入地感知环境变化与人体活动。随着特征提取与模型识别能力的不断提升，该技术的应用范畴已从早期的人体行为识别（HAR），持续向更广阔的场景拓展，展现出作为普适性环境智能（Ambient Intelligence）核心感知层的潜力。在智慧家居与养老看护领域，除了跌倒检测等典型应用外，研究已延伸至呼吸、心率等生命体征的监测，以及日常活动模式的分析，为实现健康预警与个性化服务提供数据支持。在人机交互与身份认证领域，无线信号可用于识别手势指令、构建无形的数字界面，甚至通过个体独特的动作习惯或体型反射特征实现身份识别。此外，在工业物联网、安防监测等场景中，该技术也被用于设备的无接触状态监控、入侵检测等任务。这些拓展性应用共同面临并推动了若干基础性挑战的解决：其一，是对更微弱信号（如胸腔起伏）的感知能力；其二，是在更复杂环境（多人员、多设备干扰）下的鲁棒性；其三，是对用户隐私保护更高标准的满足。应对这些挑战，不仅需要改进硬件与信号处理方案，也对后端识别模型的泛化能力、抗噪性和可解释性提出了更高要求。尽管应用场景各异，但其技术内核共享着通过无线信号扰动解析物理世界变化的同一范式。因此，在核心行为识别任务上取得的模型进步——例如对细粒度动作更强的区分能力、对噪声更稳健的特征融合机制——其价值并不仅限于动作分类本身，其设计思路（如双流特征处理、非线性融合等）亦能为上述更广泛场景中的关键难题提供有益的解决思路。本研究在Wi-Fi HAR上对相位-振幅融合、时空协同建模及非线性表征的探索，也是对这一普适性技术路线的一次深化与实践。

# 模 型

模型如图1所示，首先将原始CSI拆分提取为振幅与相位两路，其中相位先解缠去除跳变，再与归一化后的振幅特征并行输入网络，此阶段建立振幅与相位的共同表征空间，避免后期融合带来的信息损失。

在时间和通道维度上，本文采用GRE编码器与MCAT模块进行协同建模。GRE刻画了动作片段的时间跨度与相邻片段的相对次序，为后续注意力机制提供稳定的时间先验；MCAT则结合自注意力与卷积，捕捉长短程依赖与局部节奏。此时间表达能稳健地捕捉相似动作中的细微差异，提升模型对细粒度时序变体的识别能力。

在特征通道维度，本文引入可学习位置编码（LPE）与MRPT模块。LPE使网络自适应学习通道间的相对关系，避免固定位置编码带来的约束[14]。MRPT结合自注意力与卷积提炼通道间的相关性，通过残差与归一化稳定训练，充分聚合跨通道的互补信息，从而提升对细粒度差异的敏感度[15]。

在两条通道分支的末端，使用深度可分离卷积(DWConv)进行轻量级特征融合。DWConv首先独立提取局部模式，再在极少量参数下完成跨通道的信息交互[16]。提纯后的振幅与相位特征通过KAN-GRN模块进行非线性融合。该模块以门控残差为骨架，侧重建模振幅与相位之间可能存在的非线性耦合关系；其中的KAN[17]用于以一组可学习基函数对复杂映射进行展开，门控单元根据输入上下文自适应[18]调节各路信息的贡献，最终输出与分类器对接的判别表示。整体上，这一串联式改造在不破坏原有的单时间通道特征信息优势的前提下，首先振幅与相位[19]并行进入模型，其次LPE与MRPT强化通道特征信息后DWConv进行轻量融合最后由KAN-GRN非线性判别这一闭环流程，使模型能够在行为识别尤其是相似行为场景中更稳定地拉开类间边界，同时保持参数规模与推理延迟在可控范围内，面向边缘/端侧部署的进一步轻量化拟在后续工作中开展。

|  |
| --- |
| 图1 PA-TCNet模型总体架构 |

## 原始CSI数据预处理

图表

AI 生成的内容可能不正确。

图2 PA-TCNet模型总体架构

为直观反映 CSI 在时间与频率上的分布，对某一子载波的 CSI 序列进行短时傅里叶变换（STFT），得到时频图如图 X 所示。图中横轴为时间，纵轴为频率，颜色表示该时频点的功率（dB）。时频图2能够同时刻画信号随时间变化的频率成分，便于观察人体活动引起的多径变化与多普勒效应在频域上的表现，为后续特征提取与行为识别提供依据。

**2.1.1 复合滤波操作**

为保证后续模型输入的稳定性与有效性，本研究对采集到的Wi-Fi CSI原始数据进行了多阶段滤波处理。

低通滤波器用于去除高于人体动作主频的高频干扰分量。本研究将截止频率设定在3–6Hz范围内，结合双向滤波避免相位畸变。该步骤保证了CSI信号整体频谱的约束，滤除了环境抖动与设备噪声带来的高频成分[20]。

其次，对低通滤波后的序列引入汉明窗平滑操作，使序列在局部时间范围内更为平滑稳定。本研究采用窗口长度的加权滑动平均，在5–31范围内选取，依据验证集效果确定最优值。该步骤进一步增强了信号的短时稳定性，同时避免了过度平滑带来的动作边沿钝化。

采用卡尔曼滤波器在状态空间中对序列进行最优估计[28]。过程噪声协方差 Q 与观测噪声协方差 R 的设置依据为：R 表示对观测值的不确定性，本文以前级低通与汉明窗滤波后的序列作为观测，R 根据该序列在短时窗口内的残差方差或噪声功率估计设定，使滤波在抑制高频抖动的同时不过度平滑动作边沿；Q 表示状态转移的不确定性，在 CSI 时序中人体动作引起的信道变化在相邻采样间具有一定连续性，Q 取与采样间隔和典型状态变化率相适应的量级，具体数值在验证集上对识别率与平滑效果折中确定。通过合理设置 Q 与 R，卡尔曼滤波能够在保证平滑性的同时，提升对遮挡、多径干扰以及瞬时突变的鲁棒性。

**2.1.2 时间对齐与滑动窗口**

在实际的Wi-Fi CSI数据采集中，不同记录的采样长度存在差异，大多数样本的长度约为10,000。若直接输入模型，不仅难以进行批处理操作，还会导致输入维度不一致，从而影响特征提取与模型训练的稳定性。本文基于时间戳对所有样本进行了对齐处理，并将其统一裁剪或缩放至最大长度2000。

在完成序列对齐后，本文进一步采用滑动窗口策略以满足在线预测的需求，并实现数据扩增。具体设置窗口长度为1000，步长为150，在对齐后的序列上进行切分。

## 幅相处理

首先对时间对齐后的复数进行极坐标分解：

该步骤与滑窗/对齐串联，可直接得到幅度与相位分量供后续处理。

**2.2.1 相位校正**

在CSI信号中，相位信息蕴含了丰富的运动特征，但其直接利用往往面临较大挑战。一方面，原始相位序列普遍存在相位跳变现象[21]，即相位值在−π与π之间周期性折返，从而导致曲线出现不连续突变；另一方面，采样点数量庞大，使得绘制的原始相位图呈现出明显的混乱与波动，不利于后续特征提取与模式识别。

为解决上述问题，本文在相位处理中引入了相位解卷绕操作，结果如图3所示。该方法通过识别相邻采样点之间的不连续突变，并在跨越±π时自动补偿相位偏移，从而将相位曲线展开为连续平滑的形式。解卷绕后的相位数据能够更真实地反映信号随时间的变化趋势，消除了周期性折返带来的伪跳变现象，使得后续滤波与特征建模更加稳定可靠。

CSI 相位通常被包裹在 [−π, π] 或 [0, 2π] 内，解卷绕旨在消除跨边界跳变、得到连续相位序列。如递推式（3）所示。其中 为第 (i) 点的原始相位，为解卷绕后的相位*，* 为前一点解卷绕结果。

考虑到Wi-Fi通信过程中存在由硬件偏移和载波频率不同步引起的线性相位漂移[22]，本文在相位解卷绕之后进一步进行了去线性漂移处理。该操作通过拟合并消除全局线性趋势，有效去除了由发射端与接收端硬件差异造成的相位偏移，从而凸显出与人体动作直接相关的相位动态变化。

设解卷绕相位序列为，采用最小二乘法在全序列上拟合线性趋势，并将该趋势项从原序列中扣除，得到去漂移后的相位:

其中参数 由最小化全体采样点的平方残差确定：

从而突出与人体动作相关的相位变化[29]。

图片包含 日程表

AI 生成的内容可能不正确。

图3 相位解卷绕图

## 双通道特征提取

在双流建模阶段，振幅与相位特征被分别输入到两条并行分支中，并在时间维度与特征维度进行差异化建模，以实现对Wi-Fi CSI多模态特征的充分挖掘。时间通道维度结合了MCAT与GRE，与传统的单纯transformer不同，MCAT(图4)不仅包含多头自注意力机制以捕捉全局依赖关系：

其中为查询矩阵；为键矩阵；为值矩阵；为键/查询维度；对每个查询在键维做行归一化。

同时还引入了多尺度卷积残差模块[23]，使得模型能够在保留全局信息的同时，强化局部动态特征的提取。相比于只使用注意力机制，MCAT在复杂动作识别场景下能够更好地兼顾长时序依赖与短时局部特征，从而提升时序特征的鲁棒性。GRE单元则通过门控机制在时间维度上实现对关键帧的自适应选择，这一设计区别于传统残差连接，能够有效抑制冗余时序信息，使模型在捕捉长时间动作动态时避免信息过载。

在特征通道维度方面，模型结合了MRPT(图3)与LPE。传统的通道注意力方法往往仅在单尺度或全局范围内进行加权，难以兼顾局部与整体的通道依赖关系。LPE通过引入相对位置矩阵：

其中X为输入特征矩阵；为加入相对位置编码后的特征矩阵;为相对位置编码；为可学习的映射函数，用于将相对距离映射为可加性偏置。 本文LPE采用可学习参数实现，即由训练得到，而非固定周期函数。

强化了通道特征的位置信息，使得CSI特征在时间域和特征域均具备更强的顺序感知能力。而MRPT在传统投影-融合机制的基础上，设计了多尺度残差投影结构，分别进行尺度级与全局级的跨通道信息整合。这一设计相较于单一注意力机制[24]，能够更好地实现不同尺度通道特征的互补性与判别性，提升模型的整体表达能力。

此外，考虑到Wi-Fi通信过程中存在由硬件偏移和载波频率不同步引起的线性相位漂移，本文在相位解卷绕之后进一步进行了去线性漂移处理。该操作通过拟合并消除全局线性趋势，有效去除了由发射端与接收端硬件差异造成的相位偏移，从而凸显出与人体动作直接相关的相位动态变化。在双流输出阶段，振幅与相位分支的特征首先经过DWConv进行降维与参数压缩：

其中为输入特征；为depthwise卷积核；为pointwise卷积核。为DWConv后的输出特征，用于降维与参数压缩。

该设计显著降低了计算开销，使模型能够在保证性能的同时在双流融合环节降低计算、有利于参数与计算控制。随后，两个模态的特征通过门控残差机制进行融合：

其中和为振幅分支与相位分支的特征；为门控映射参数；为门控偏置；为Sigmoid函数。为Hadamard乘。

在该过程中振幅与相位的互补性被充分利用，融合后的特征具备全局与局部、多模态与多尺度的综合优势。



图4 MCAT和MRPT模块结构图

与传统单一注意力模型相比，本研究的双流Transformer架构在时间通道维度采用的MCAT与GRE增强时序动态建模，在特征通道维度采用MRPT与LPE强化跨通道特征表达，最终结合DWConv与门控融合机制实现在融合阶段兼顾参数效率与鲁棒性的特征表征。这种组合注意力机制不仅弥补了单一注意力结构在局部建模与尺度建模上的不足，也显著提升了Wi-Fi CSI人体动作识别的精度与泛化能力。

**1.4特征融合层：KAN-GRN**

为增强幅相耦合的表达与稳健融合，本文在融合位点引入KAN-GRN联合单元。KAN通过将每条权重替换为可学习的一元函数，使得网络能够更加高效地逼近复杂的非线性映射，并且显著增强了网络对于高阶非线性关系的刻画能力。与传统的MLP相比，KAN不仅能够在高维空间中展开幅度与相位之间的复杂耦合模式，还能够在保持较少参数的情况下提升模型的表达能力。受KAN相关工作的启发，我们将其引入到振幅与相位特征的融合模块中，以实现幅相耦合关系的高维展开与精细判别。

具体而言，KAN的可学习基函数完成非线性展开与稀疏逼近，以刻画幅度与相位间的高阶耦合；同时，GRN（Gate Residual Network）的门控残差结构用于信息选择与稳定融合，以抑制环境扰动与冗余通道。两者协同作用，一方面提升了对细粒度相似动作的判别力，另一方面保持了训练的收敛性与计算开销的可控性。KAN-GRN结构如图5所示。

在具体实现中，设振幅支路与解缠相位支路在聚合后分别得到向量，记为拼接向量。经典GRN可写成公式12和13。

其中与为非线性与门函数，为残差。我们将其中的线性映射替换为KANLinear：

其中为一组可学习基函数，为可学习系数。其中 为一组可学习基函数， 为可学习系数。本文中 采用小波基实现：设第 条连接上的尺度与平移为可学习参数 ，则 ，取 Mexican Hat 小波 或 Morlet 小波等，通过可学习尺度与平移自适应刻画幅相非线性耦合。初始化方面线性项对应权重采用 Kaiming 均匀初始化，小波系数 在给定尺度下均匀初始化，尺度与平移 初值分别取 1 与 0。由此得到公式15：

KAN负责将线性可分性不强的幅度与相位耦合关系展开到高维可分空间，从而增强了特征之间的可分性；而GRN则负责在噪声与冗余通道的干扰下进行选择性信息传递与幅度抑制，保证信息融合的稳定性。残差连接则保证了梯度流畅与原始判别信号不被破坏。



图5 KAN结构图

与仅用线性门控[25]或MLP的融合方式相比，KAN-GRN协同机制显著提升了整体性能。KAN的可学习多尺度基函数在高维空间精细刻画了幅度变化与相位漂移的耦合模式，从而增强了细粒度特征的可分性；同时，GRN的逐通道门控在多径扰动下抑制了不稳定的子空间，保持了融合结果的统计稳态，减少了误报与漏检。结合残差与层归一化，保证了数值与优化的稳定性，并使得该机制能够与模型主干无缝衔接，从而在几乎不增加时延的前提下持续提升精度。

最后，KAN的引入不仅增强了模型对幅度与相位之间复杂耦合关系的捕捉能力[26]，还能有效减少噪声与环境变化对行为识别性能的负面影响。通过自适应地展开高阶特征映射，KAN使得模型能够在复杂的环境中保持高识别精度。与传统的特征融合方法[27]相比，KAN-GRN通过更高效的特征表示与选择性信息融合，不仅提升了模型的稳定性，还在在融合模块层面具有较好参数–性能比、有利于后续剪枝/量化等扩展。尤其在细粒度动作识别与跨环境迁移任务中，KAN-GRN模块展现了优越的鲁棒性和适应性。

# 实验设置与分析

## 数据集

为全面评估模型在日常行为动作与相似动作的区分能力，本研究选取两套公共数据集(StanWiFi和MultiEnv)和一套自建数据集(SimAct)，且数据集均包含振幅与相位分量，数据集设置如表1所示。

**表1 数据集(StanWiFi/****MultiEnv/SimAct)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 数据集 | 场景 | 设备配置 | 类别数 | 类别集合 | 受试者数 | 总样本数 |
| StanWiFi | General | Tx=1, Rx=3 | 6 | 跌倒、跑步、躺下、走路、坐下、站立 | 6 | 577 |
| MultiEnv-办公室 | Office | Tx=1, Rx=3 | 6 | 静止、跌倒、坐下、起立、行走、转身、拾取 | 30 | 1000 |
| MultiEnv-大厅 | Hall | Tx=1, Rx=3 | 6 | 静止、跌倒、坐下、起立、行走、转身、拾取 | 30 | 1000 |
| SimAct | Lab | Tx=2, Rx=2 | 6 | 挥手、招手、推、拉、坐下、蹲下 | 10 | 2000 |

**2.1.1 SimAct**

图形用户界面, 应用程序

AI 生成的内容可能不正确。

图6 实验场景布局

自建数据集SimAct主要面向相似动作的细粒度识别，在会议室场景采集了由10名受试者执行的6类动作(挥手、招手、坐下、蹲下、推和拉)，实验场景布局如图6所示。采集系统基于Intel 5300无线网卡，采样率为1000 Hz，收发机之间距离2m。

## 实验参数设置

本项工作在NVIDIA GeForce RTX 4070Ti Super显卡上完成模型训练与验证，运行环境为CUDA 12.4、Python 3.11，并采用TensorFlow/PyTorch深度学习框架实现。

## 实验结果

**2.3.1 StanWiFi实验结果**

表2展示了本文模型与基准模型和最新模型在StanWiFi数据集上进行比较分析的结果，采用的评估标准包括准确率(Acc)、精确度(Pre)、召回率和F1-Score。表中部分对比模型（如 LSTM、THAT、RGANet）的 Pre、Recall 或 F1 项标注为“—”，系依据原论文及公开代码复现时，所用实现未提供上述指标的输出，故无法在本实验中一并给出；表中已对缺省项在表注中予以说明。在单场景StanWiFi数据集识别下，本文模型四项指标同时最优且识别率达到99.83%，核心在于表征能力与融合策略的提升：LPE通过数据驱动的时间对齐，提高了对不同动作节律与停顿的敏感度；MRPT在通道维度引入相对关系约束，增强了子载波/天线间的可分性；KAN-GRN以非线性基函数实现幅度与相位的高阶耦合，使类间间隔更大、决策边界更平滑。进一步地，在同等参数量下，本模型的有效表征容量更高，训练过程收敛更快且振荡更小，减少了由优化不稳导致的性能波动。

**表2 StanWiFi数据集上的实验结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | Acc | Pre | Recall | F1-Score |
| LSTM | 91.4 | - | - | - |
| THAT | 97.2 | - | - | - |
| CNN-GRU | 95 | 94.23 | 94.65 | 95.01 |
| RGANet | 97.31 | 99.5 | 99.43 | - |
| 本文模型 | 99.83 | 99.86 | 99.63 | 99.75 |

**2.3.2 MultiEnv实验结果**

**表3 MultiEnv数据集上的实验结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验环境 | Model | Acc | Pre | Recall | F1-Score |
| 办公室 | LSTM | 94.03 | - | - | - |
| THAT | 96.95 | 96.28 | 96.26 | 97.26 |
| CNN-GRU | 97 | 98.16 | 97.87 | 97.01 |
| RGANet | 97.31 | 98.5 | 98.43 | - |
| 本文模型 | 99.13 | 99.28 | 99.16 | 98.98 |
| 大厅 | LSTM | 91.03 | - | - | - |
| THAT | 96.45 | 96.15 | 96.31 | 96.12 |
| CNN-GRU | 97 | 97.53 | 97.23 | 98.76 |
| RGANet | 97.31 | 98.5 | 98.46 | - |
| 本文模型 | 98.67 | 98.56 | 98.9 | 98.78 |

表3汇总了在两个典型场景(办公室与大厅)下，各模型在MultiEnv数据集上的表现。在办公室与大厅两种环境下，多径结构与噪声分布差异显著。本文模型之所以在两表均为最优，源于LPE缓解不同环境造成的时序偏移与频率失衡；MRPT通过相对位置信息建模抑制频率选择性衰落带来的通道扰动；DWConv以轻量多尺度卷积覆盖快慢节律与长短依赖，在控制参数量的同时降低过拟合倾向；KAN-GRN在分布变化下仍能稳健整合幅相信息，提升边界的可分与稳定性。实证上，该设计对设备摆放、人员分布与遮挡情况的小幅变化不敏感，从而在不同环境条件下均取得更高精度与更稳定的表现。

**2.3.3 SimAct实验结果**

表4比较了THAT模型和本文模型在面向相似动作的自制数据集SimAct上的性能。可以发现本文模型性能远超THAT模型，在所有评估的指标中均达到了95.1%，进一步证明了本模型在相近动作上的判别优势。

**表4 SimAct数据集上的实验结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Acc | Pre | Recall | F1-Score |
| THAT | 91.2 | 91.2 | 91.2 | 91.2 |
| 本文模型 | 95.1 | 95.1 | 95.1 | 95.1 |

## 消融实验

为评估各模块贡献，本文进行了消融实验(表5)。整体上，移除任一模块都会使识别率下降，但不同场景的敏感度不同。LPE对时序可分性最关键：去除后四个场景的识别率均下降，其中办公室场景跌幅最大，表明LPE能有效刻画短促节律与相邻停顿差异，从而减少相似动作混淆。相对位置注意力在复杂多径与相似动作下带来稳定提升：StanWiFi与SimAct对其更为敏感，该机制可聚焦关键子载波与关键时刻并抑制环境干扰。MRPT主要负责跨通道的相对依赖与对齐：移除后StanWiFi与MultiEnv-大厅识别率分别下降约2.3与1.3个百分点，MultiEnv-办公室变化较小，而SimAct的下降最为明显。

DWConv在几乎不影响识别率的前提下显著降低计算与参数；KAN-GRN明显提升相似动作的可分性，在StanWiFi与MultiEnv-办公室数据集上提升更突出，在MultiEnv-大厅与SimAct数据集上同样带来提升并有效拉开类别边界。仅用振幅会显著削弱性能，说明相位在方向性、微小位移与加减速节律等细粒度差异上不可替代。综上，LPE与KAN-GRN分别奠定了时序定位与高层判别的上限，相对位置注意力增强复杂环境下的鲁棒性，DWConv实现精度与效率平衡，相位提供关键补充；多模块协同，使完整模型在各评测场景均取得最优表现。表 5 同时给出将 KAN-GRN 替换为结构相同的 MLP-GRN 的对照。在四类场景下，KAN-GRN 均优于 MLP-GRN，表明 KAN 在幅相融合上的增益来自其可学习基函数带来的非线性表达能力，而非单纯门控结构。

**表5消融实验**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Stan | | MultiEnv office | | MultiEnv Hall | | SimAct | |
| Accuracy | Δ | Accuracy | Δ | Accuracy | Δ | Accuracy | Δ |
| 本文模型 | 98.1 | - | 97.8 | - | 98.6 | - | 95.1 | - |
| w/o LPE | 93.9 | -4.1 | 91.2 | -6.5 | 90.9 | -7.6 | 88.5 | -6.5 |
| w/o MRPT | 95.8 | -2.2 | 97.6 | -0.2 | 97.3 | -1.2 | 91.5 | -3.6 |
| w/o DWConv | 97.4 | -0.8 | 97.3 | -0.5 | 98.2 | -0.4 | 92.9 | -2.3 |
| w/o KAN-GRN | 93.8 | -4.4 | 95.8 | -1.9 | 97.7 | -0.9 | 93.2 | -1.9 |
| KAN**→MLP (**MLP-GRN) | 96.2 | -1.9 | 96.9 | -0.9 | 98.0 | -0.6 | 94.2 | -0.9 |
| Amplitude Only | 92.1 | -5.9 | 89.8 | -7.9 | 92.2 | -6.4 | 90.1 | -4.9 |

# 结论

本研究提出的PA-TCNet双通道模型在人体行为识别尤其是相似行为中表现优异。以双分支注意—多尺度卷积编码器为主干，结合相位与振幅特征、可学习位置编码(LPE)与通道相对注意力MRPT的表征强化，并以KAN-GRN实现非线性融合。实验表明，相位信息的加入提升了顺序敏感与方向相反类别的稳定性；LPE明显增强跨速度与序列长度下的时序可分性；MRPT强化了通道间的相对关系建模，尤其在短促动作与方向性差异的细粒度区分上更为有效；KAN-GRN在融合层带来最为稳定的判别增益。由此，模型在多场景下均实现了多评价指标的同步改善，验证了整体设计路径对人体行为识别尤其是相似行为识别具有系统性成效。

未来我们将面向跨环境场景扩展模型能力，例如引入迁移学习以稳健应对设备与场地变化、提升动作区分度。其次，多模态融合也是提升模型性能的重要方向。面向智慧医疗、智能家居等资源受限场景，将开展模型轻量化、剪枝/蒸馏/量化及边缘部署与实时性验证，推动方法从实验室向实际应用过渡。除了现有的Wi-Fi CSI数据，未来我们将考虑引入其他传感器数据，如视觉和声音等，以进一步提升模型的行为识别能力，从而服务跌倒预警与康复评估等医疗健康应用，助力居家照护与智慧医疗发展。

# 参 考 文 献:

1. 陈莹,龚苏明.改进通道注意力机制下的人体行为识别网络[J].电子与信息学报,2021,43(12):3538-3545.
2. Alsaify B A, Almazari M, Alazrai R, Alouneh S, Daoud M I. A CSI-Based Multi-Environment Human Activity Recognition Framework [J]. Appl. Sci., 2022, 12(2):930.
3. Varga D. Mitigating Data Leakage in a WiFi CSI Benchmark for Human Action Recognition [J]. Sensors, 2024, 24(24):8201.
4. Abuhoureyah F, Sim K S, Wong Y C. Multi-user human activity recognition through adaptive location-independent WiFi signal characteristics [J]. IEEE Access, 2024, 12:112008-112024.
5. Wang F, Gong W, Liu J. On spatial diversity in WiFi-based human activity recognition: A deep learning-based approach[J]. IEEE Internet of Things Journal, 2018, 6(2): 2035-2047.
6. Chen Z, Zhang L, Jiang C, Cao Z, Cui W. WiFi CSI Based Passive Human Activity Recognition Using Attention Based BLSTM [J]. IEEE Trans. Mob. Comput., 2019, 18:2714-2724.
7. Moshiri F P, Shahbazian R, Nabati M, Ghorashi S A. A CSI-Based Human Activity Recognition Using Deep Learning [J]. Sensors, 2021, 21(21):7225.
8. Shi Z, Cheng Q, Zhang J A, et al. Environment-robust WiFi-based human activity recognition using enhanced CSI and deep learning[J]. IEEE Internet of Things Journal, 2022, 9(24): 24643-24654.
9. Deng F, Jovanov E, Song H, et al. WiLDAR: WiFi signal-based lightweight deep learning model for human activity recognition[J]. IEEE Internet of Things Journal, 2023, 11(2): 2899-2908.
10. Yang J, Chen X, Zou H, Lu C X, Wang D, Sun S, Xie L. SenseFi: A Library and Benchmark on Deep-Learning-Empowered WiFi Human Sensing [J]. Patterns, 2023, 4(3):100703.
11. Li B, Cui W, Wang W, Zhang L, Chen Z, Wu M. Two-Stream Convolution Augmented Transformer for Human Activity Recognition [J]. Proc. AAAI Conf. Artif. Intell., 2021, 35(1):286-293.
12. Luo F, Khan S, Jiang B, et al. Vision transformers for human activity recognition using WiFi channel state information[J]. IEEE Internet of Things Journal, 2024, 11(17): 28111-28122.
13. 朱煜,赵江坤,王逸宁,等.基于深度学习的人体行为识别算法综述[J].自动化学报,2016,42(06):848-857.DOI:10.16383/j.aas.2016.c150710.
14. Gao Q, Wang J, Ma X, Feng X, Wang H. CSI-based device-free wireless localization and activity recognition using radio image features [J]. IEEE Transactions on Vehicular Technology, 2017, 66 (11):10346-10356.
15. Hussain A, Hussain T, Ullah W, et al. Vision transformer and deep sequence learning for human activity recognition in surveillance videos [J]. Computational Intelligence and Neuroscience, 2022, 2022:3454167.
16. Yadav S K, Sai S, Gundewar A, et al. CSITime: Privacy-preserving human activity recognition using WiFi channel state information [J]. Neural Networks, 2022, 146:11-21.
17. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. Advances in neural information processing systems, 2017, 30.
18. Ding X, Jiang T, Zhong Y, et al. Wi-Fi-based location-independent human activity recognition with attention mechanism enhanced method [J]. Electronics, 2022, 11 (4):642.
19. Xia K, Huang J, Wang H. LSTM-CNN architecture for human activity recognition [J]. IEEE Access, 2020, 8:56855-56866.
20. Mutegeki R, Han D S. A CNN-LSTM approach to human activity recognition[C]//2020 international conference on artificial intelligence in information and communication (ICAIIC). IEEE, 2020: 362-366.
21. Vrskova R, Kamencay P, Hudec R, et al. A new deep-learning method for human activity recognition [J]. Sensors, 2023, 23 (5):2816.
22. Khatun M A, Yousuf M A, Ahmed S, et al. Deep CNN-LSTM with self-attention model for human activity recognition using wearable sensor [J]. IEEE Journal of Translational Engineering in Health and Medicine, 2022, 10:1-16.
23. Khatun M A, Yousuf M A, Ahmed S, et al. Deep CNN-LSTM with self-attention model for human activity recognition using wearable sensor [J]. IEEE Journal of Translational Engineering in Health and Medicine, 2022, 10:1-16.
24. Liu Z, Wang Y, Vaidya S, et al. Kan: Kolmogorov-arnold networks[J]. arXiv preprint arXiv:2404.19756, 2024.
25. Vaca-Rubio C J, Blanco L, Pereira R, et al. Kolmogorov-arnold networks (kans) for time series analysis[J]. arXiv preprint arXiv:2405.08790, 2024.
26. Bodner A D, Tepsich A S, Spolski J N, et al. Convolutional kolmogorov-arnold networks[J]. arXiv preprint arXiv:2406.13155, 2024.
27. Tokunaga K, Xiang Y, Nagazawa R, et al. Pre-processing of CSI signal for Wi-Fi sensing-based motion detection[J]. Nonlinear Theory and Its Applications, IEICE, 2025, 16(3): 506-518.
28. Li C, Brauer J, Sezgin A, et al. Kalman filter based MIMO CSI phase recovery for COTS WiFi devices[C]//ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Toronto, ON: IEEE, 2021: 4985-4989.
29. Quy T D, Lin C Y, Shih T K. Enhanced human activity recognition using wi-fi sensing: Leveraging phase and amplitude with attention mechanisms[J]. Sensors, 2025, 25(4): 1038.