

基于多尺度特征融合的跨视角点云步态识别^{*}魏永超¹ 谢卫鑫² 张姝岚² 王应海² 孙如新²

(1. 中国民用航空飞行学院科研处 德阳 618307; 2. 中国民用航空飞行学院计算机学院 德阳 618307)

摘 要: 现有的步态识别方法大多基于剪影或骨骼, 然而二维信息缺乏对人体空间几何结构的完整描述, 在视角变化、遮挡等复杂条件下识别效果性能有限, 为此本文提出了一种结合全局多尺度和局部细粒度特征的点云步态识别方法。该方法将点云投影为深度步态图, 引入跨视角数据变换模块提升模型的视角不变性, 采用改进的残差网络提取丰富的全局多尺度步态特征, 最后使用 KAN 网络增强局部细粒度步态特征的代表力。实验结果表明, 基于点云的步态识别方法远优于基于二维信息的方法, 该方法在 SUSTech1K 公开数据集上取得了 92.65% 的平均 Rank1 准确率, 相较于先进方法 LidarGait 提升了 6.02%, 充分验证了该方法的有效性。

关键词: 步态识别; 残差网络; KAN 网络; 数据变换; 视角不变性

中图分类号: TN249 **文献标识码:** A **国家标准学科分类代码:** 510.20

Cross-view point cloud gait recognition based on multi-scale feature fusion

Wei Yongchao¹ Xie Weixin² Zhang Yalan² Wang Yinghai² Sun Ruxin²

(1. Scientific Research Office, Civil Aviation Flight University of China, Deyang 618307, China;

2. School of Computer Science, Civil Aviation Flight University of China, Deyang 618307, China)

Abstract: Most of the existing gait recognition methods are based on silhouettes or skeletons, however, the 2D information lacks a complete description of the spatial geometry of the human body, and the performance of the recognition effect is limited under complex conditions such as view angle change and occlusion, for this reason, this paper proposes a point cloud gait recognition method that combines global multiscale and local fine-grained features. The method projects the point cloud as a depth gait map, introduces a cross-view data transformation module to improve the viewpoint invariance of the model, uses an improved residual network to extract rich global multi-scale gait features, and finally uses a KAN network to enhance the representativeness of local fine-grained gait features. The experimental results show that the gait recognition method based on point cloud is far better than the method based on 2D information, which achieves an average Rank1 accuracy of 92.65% on the SUSTech1K public dataset, which is a 6.02% improvement compared to the advanced method LidarGait, which fully verifies the effectiveness of the method.

Keywords: gait recognition; residual network; KAN network; data transformation; viewpoint invariance

0 引 言

步态识别作为一种新兴的生物特征识别技术, 通过利用个体行走模式的独特性进行身份验证。与指纹、人面、虹膜等传统生物特征相比, 步态识别具有无接触、远距离识别的独特优势^[1], 在社会安全、刑侦追踪^[2]等领域具有重要的应用价值。然而, 步态识别在实际应用中仍面临许多挑战, 尤其是同一目标的步态因视角变化、遮挡、穿着变化和携带物品等协变量因素而产生显著外观差异, 增加模型学习一致性特征的难度。

近年来, 基于二维信息的步态识别方法^[3-8]取得了较大进展, 这些方法依赖于摄像头采集步态序列, 通过提取轮廓^[9]、骨骼^[10]等特征实现身份识别。然而, 二维信息缺乏对人体空间几何结构的完整描述, 且成像容易受到光照等条件的影响^[11]。因此, 不少研究者开始转向基于三维信息步态识别研究^[12-13]。基于点云的步态识别尚属起步阶段, 主要有两种研究方向, 第 1 种是直接点云提取步态特征, 第 2 种是将点云投影为二维点图, 进而使用卷积神经网络提取步态特征。该领域已有一些重要进展, Shen 等^[14]提出了一种基于点云的步态识别框架 LidarGait, 该方法通过将

稀疏点云投影至深度图像,结合残差网络提取全局步态特征。Han 等^[15]提出了 HMRNet 网络,采用双分支结构分别从点云的几何特征和距离视图的结构特征中进行学习。然而,尽管这些方法取得一定成果,但现有的研究主要集中在全局步态特征的提取,忽略了局部细粒度特征的建模。此外,点云在提供丰富人体空间信息的同时,仍受视角变化、遮挡等因素影响,现有方法并未显式地解决这一问题,而是通过卷积等操作隐式地提取一致性特征。为了解决上述问题,本文提出一种新的步态识别框架,主要贡献可以概括如下:

1) 考虑到步态识别是细粒度特征学习的任务,本文提出一种结合全局多尺度特征和局部细粒度特征的步态识别模型结构。在 SUSTech1K 公开数据集上的大量实验表明,本文方法与当前先进方法相比更具竞争力。

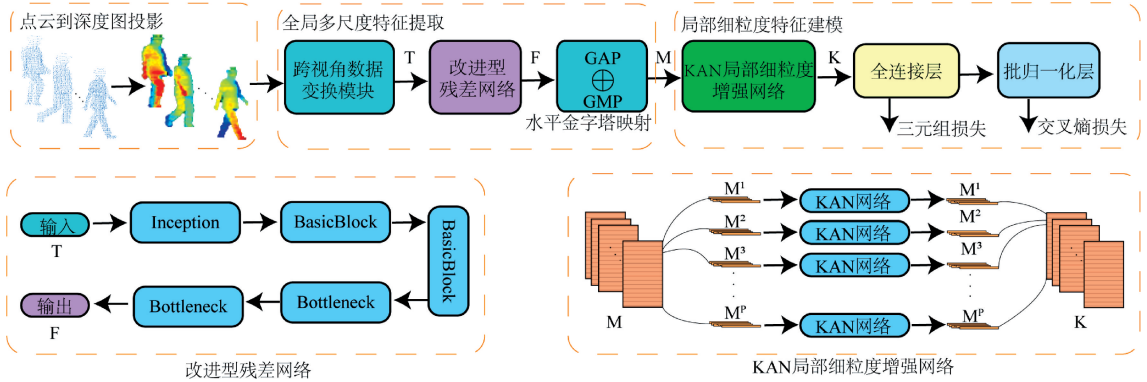


图 1 基于多尺度特征融合的跨视角点云步态识别整体框架

Fig. 1 Holistic framework for Cross-view point cloud gait recognition based on multi-scale feature fusion.

首先,将点云步态数据从 3D 空间投影到 2D 平面,通过颜色编码将二维点图映射为深度步态图,在保留高精度三维人体结构信息的同时,显著降低后续模型计算复杂度。本文将处理后的深度步态图作为训练数据。

输入 n 帧的深度步态图 $\{I_j \in R^{h \times w}, j = 1, 2, \dots, n\}$, 每帧高为 h , 宽为 w 。输入的原始数据 I 首先经过跨视角数据变换模块 $CPT(\cdot)$, 模拟不同视角的步态数据, 丰富步态样本的多样性:

$$T = CPT(I) \quad (1)$$

式中: $T \in R^{n \times c_1 \times h \times w}$, h 和 w 保存不变, c_1 为通道数。之后将变换数据 T 送入一个具有 5 层结构的改进型残差网络 $IRN(\cdot)$, 提取全局多尺度步态特征:

$$F = IRN(T) \quad (2)$$

式中: $F \in R^{n \times c_1 \times h_2 \times w_2}$, c_1 , h_2 和 w_2 分别是主干网络提取后的特征 F 通道数及空间尺寸。 IRN 网络由一个初始卷积和 4 层残差结构堆叠而成, 之后通过时间池化模块将不定长的步态序列沿时间维度最大池化聚合步态特征, 生成更加紧凑且具有代表性的帧级特征。之后将帧级特征送入 HPP 模块进行水平金字塔映射, 该模块主要是将帧级特征从水平上划分 p 个区域, 并对每个帧级子区域进行全局平

2) 本文设计了跨视角数据变换模块, 通过模拟步态序列的视角变化, 提升模型视角不变性, 增强了模型鲁棒性。

3) 设计改进型残差网络用于提取全局多尺度步态特征, 引入 KAN 网络增强局部细粒度步态特征的表征力, 通过对局部区域精细化建模, 提升步态特征判别力。

1 基于多尺度特征融合的跨视角点云步态识别方法

1.1 点云步态识别框架

本文提出了一种结合全局多尺度特征和局部细粒度特征的点云步态识别方法, 将点云投影为深度步态图, 使用卷积神经网络提取步态特征。如图 1 所示, 该模型主要由跨视角数据变换模块、改进型残差网络、KAN 局部细粒度增强网络 3 个部分组成。

均池化 $GAP(\cdot)$ 和全局最大池化 $GMP(\cdot)$ 操作, 生成空间层次化特征 $M \in R^{n \times c_1 \times p}$:

$$M = GAP(F) + GMP(F) \quad (3)$$

经过 HPP 模块后, 帧级特征被划分为 p 个局部区域, 将每个局部区域单独送入 KAN 网络^[16] 增强局部细粒度特征, 再将增强后的局部细粒度特征堆叠为原始结构:

$$K = KAN(M) \quad (4)$$

式中: $K \in R^{n \times c_2 \times p}$, c_2 和 p 分别是通道数和局部区域数。KAN 网络通过 Base-样条插值和自适应加权机制进行精细化建模, 增强步态数据中的细微局部变化, 提升局部特征的判别力。最后使用全连接层和批归一化层对特征进一步映射和调整, 使用三元组损失和交叉熵损失监督训练过程。1.2~1.4 节将分别对本文提出的跨视角数据变换模块、改进型残差网络、KAN 局部细粒度增强网络进行详细介绍。

1.2 跨视角数据变换模块

步态识别面临的核心挑战之一是步态数据的复杂性。如图 2 所示, 同一目标的步态特征因穿着、携带物品、遮挡以及视角变化等因素而显著变化, 尤其是视角变化导致的

步态外观轮廓差异,会显著增加模型学习一致性步态特征的难度。

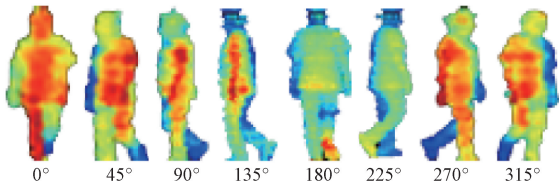


图2 同一样本的不同视角外观

Fig.2 Appearance of the same sample from different viewpoints

针对这一问题,本文提出跨视角数据变换模块,以提升模型的视角不变性与对步态协变量的鲁棒性,变换效果如图3所示。

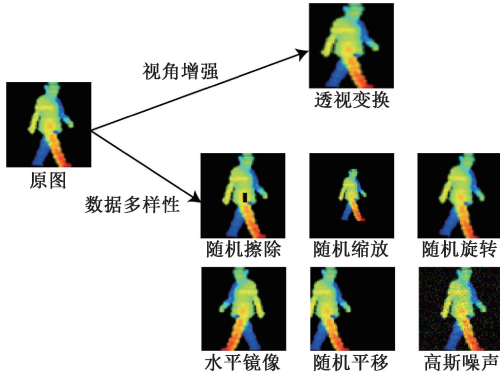


图3 跨视角数据变换效果

Fig.3 Cross-view data transformation effect

跨视角数据变换模块包括视角增强模块和数据多样性模块。从图2可看出,视角变化会显著影响步态的几何形态,为此,视觉增强模块基于透视变换策略模拟视角变化引起的步态几何变形,生成仿真视角的步态数据。通过这种方式增强模型的跨视角适应能力,提升模型的视角不变性。

数据多样性模块采用6种多样性策略:随机擦除通过遮挡局部区域模拟步态信息丢失;随机缩放通过调整目标比例模拟不同拍摄距离;随机旋转通过设定旋转角度应对运动轨迹角度变化;水平镜像生成左右对称的步态图像,增强特征学习的一致性;随机平移通过位移图像位置应对目标水平偏移;随机噪声通过添加高斯噪声提升模型对干扰的适应性。这些策略通过联合作用,缓解了因遮挡、尺度变化、视角偏移及噪声干扰等问题产生的影响,丰富了步态数据的多样性,使模型在复杂场景下具备更强的泛化能力和鲁棒性。

1.3 改进型残差网络

1)基础残差网络

ResNet^[17]自提出以来广泛应用于目标检测^[18]、语义分割^[19]等计算机视觉任务,其核心通过引入残差连接缓解深层网络的梯度消失问题,提升网络的训练效率和性能。当前,步态识别研究大多是通过堆叠多个标准卷积或设计

一种新型卷积模块构建主干网络。本文的基础残差网络是基于ResNet经典架构的轻量化残差网络,详细结构如表1所示。网络由一个初始卷积和四层标准残差块(BasicBlock)堆叠而成,逐层提取从浅层到深层的特征信息。单纯使用基础残差网络,则模型退化为LidarGait。

表1 基础残差网络结构表

Table 1 Basic residual network structure table

Layer	Structure $[k, k, c] \times b$
Conv0	$[3 \times 3, 64] \times 1$, stride=1
BasicBlock1	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 1$, stride=1
BasicBlock2	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 128 \end{bmatrix} \times 1$, stride=2
BasicBlock3	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 256 \end{bmatrix} \times 1$, stride=2
BasicBlock4	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 512 \end{bmatrix} \times 1$, stride=1

2)改进型残差网络

与基础残差网络相比,改进型残差网络有3处改进:(1)初始卷积引入Inception模块,实现了全局多尺度特征提取;(2)从单一残差结构到混合残差结构,优化了模型的特征表达能力;(3)使用ReLU6激活函数,提升了模型训练稳定性,具体结构如图4所示。

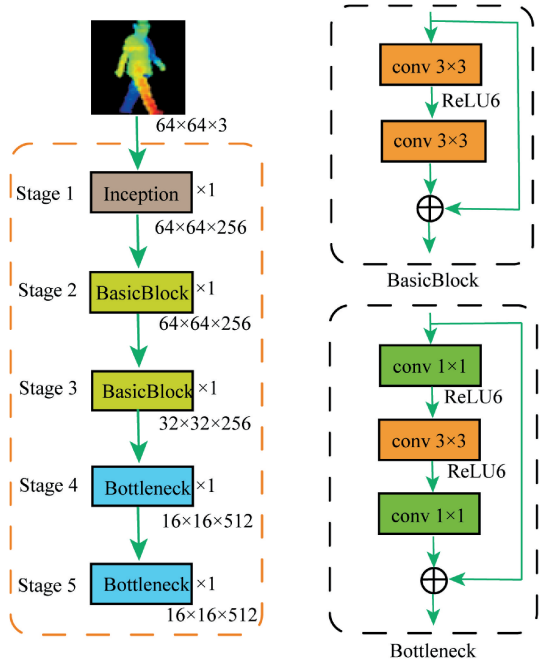


图4 改进型残差网络结构

Fig.4 Improved residual network structure

(1)Inception 模块

Inception 模块参考 InceptionV2^[20]结构,通过4个并

行分支处理不同感受野的信息,可以有效提取全局多尺度步态特征,具体结构如图 5 所示。

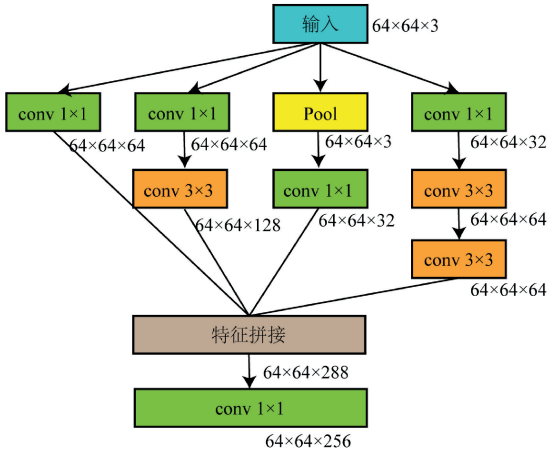


图 5 Inception 模块结构

Fig. 5 Inception module structure

步态的细节特征往往体现在局部区域,而全局特征则反映了步态的整体运动模式。Inception 模块通过并行的 1×1 、 3×3 卷积和池化操作,同时捕捉局部细节特征、中尺度特征和全局特征,提升步态特征在不同视角、遮挡等复杂因素时的判别力。最后,拼接各个分支输出的特征,形成全局多尺度步态特征,并通过 1×1 卷积进行通道整合,形成 256 维的特征输出。

(2) 混合残差网络

瓶颈残差块(Bottleneck)由 1×1 降维卷积、 3×3 中间卷积和 1×1 升维卷积组成。在 ResNet 中,BasicBlock 用于构建浅层网络,适合提取边缘信息和局部细节等高分辨率特征;Bottleneck 用于构建深层网络,通过通道压缩与扩展降低计算量的同时聚合更深层的语义信息。本文使用的点云数据拥有丰富的三维人体结构信息,步态识别不仅需要提取细粒度的轮廓信息,更需要提取深层语义信息。为此,在网络前段采用两层 BasicBlock 提取浅层特征,保证高分辨特征不被过早破坏,在中后段采用两层 Bottleneck,降低计算量的同时提高深层语义特征的表达力。

(3) ReLU6 激活函数

ReLU6 激活函数的定义为:

$$f(x) = \min(\max(x, 0), 6) \quad (5)$$

由定义可知,ReLU6 的输出范围为 $[0, 6]$,而 ReLU 的输出范围为 $[0, \infty]$,这种范围压缩使得网络更加关注小幅度的特征变化,保留弱特征信息,提升了捕捉细微动作变化的灵敏度。

1.4 KAN 局部细粒度增强网络

KAN 局部细粒度增强网络由 p 个并行的 KAN 网络组成。从 1.1 节中的描述可知,步态特征经过 HPP 模块后,帧级特征被划分为 p 个局部区域,每个局部区域被单

独送入权重不共享的 KAN 网络进行精细化建模,增强局部细粒度特征的判别力。下面将详细介绍 KAN 网络的结构。KAN 是一种受 Kolmogorov-Arnold 定理启发的新型神经网络架构,具体结构如图 6 所示。

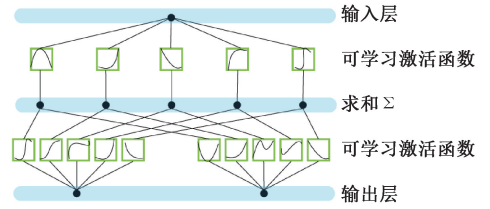


图 6 KAN 网络结构

Fig. 6 KAN network structure

多层感知机(MLP)通常在网络边缘上设置权重参数,在神经元上设置固定的激活函数。与 MLP 不同的是,KAN 将权重参数替换为被参数化为样条的单变量函数,使得网络边缘设置为可学习的激活函数,并使用自适应的方法调整模型的学习过程。从神经网络学习形式上看,KAN 网络实现从“可学习权重+固定激活函数”到“可学习激活函数+函数求和”的转变。KAN 网络的数学表达式为:

$$KAN(x) = f(x_1, \dots, x_n) = \sum_{q=1}^{2n+1} \phi_q \left(\sum_{p=1}^n \phi_{q,p}(x_p) \right) \quad (6)$$

式中: x 是 n 维的输入向量, $\phi_{q,p}: [0, 1] \rightarrow \mathbb{R}$ 是被参数化为样条的单变量函数, $\phi_{q,p}(x_p)$ 是可学习的激活函数, $\phi_q: \mathbb{R} \rightarrow \mathbb{R}$ 。可学习的激活函数 $\phi(x)$ 由基础函数 $b(x)$ 和样条函数 $spline(x)$ 组成:

$$\phi(x) = w_b b(x) + w_s spline(x) \quad (7)$$

$$b(x) = silu(x) = x / (1 + e^{-x}) \quad (8)$$

$$spline(x) = \sum_i c_i B_i(x) \quad (9)$$

式中: w_b 和 w_s 是参数化权重, $b(x)$ 是 silu 激活函数, $B_i(x)$ 是 Base-样条函数。在训练过程中,根据输入特征的分布情况,利用自适应网格提高样条函数对特征分布的拟合能力。KAN 网络通过基础线性变换进行全局特征表示,Base-样条函数捕获局部非线性模式,自适应网格优化特征分布感知能力,确保在关键区域进行细粒度特征增强。

2 实 验

2.1 数据集与预处理

SUSTech1K 数据集是一个大规模多模态步态识别数据集,总样本数为 1 050 人,包含 RGB、剪影和点云 3 种数据类型,其中点云数据是第一个基于 LiDAR 的步态数据,对于研究三维步态识别具有重要意义。如图 7 所示,SUSTech1K 除了包含主流步态数据常见的正常(normal, NM)、背包(bag, BG)、衣服(clothing, CL)、携带

(carrying, CR) 的情况,还考虑了实际生活场景雨伞(umbrellas, UB)、制服(uniform, UF)、遮挡(occlusion, OC)和夜晚(night, NT)的情况。每种步态协变量均采集 12 个视角的步态数据,分别是 0°、0°far、45°、90°、90°far、135°、180°、180°far、225°、270°、270°far、315°。

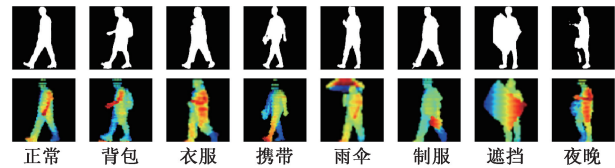


图 7 不同步态协变量对比图

Fig. 7 Comparison of different gait covariates

本文采用 SUSTech1K 的点云数据作为研究对象,并将点云数据从 3D 空间投影到 2D 平面,生成前视深度图像。在横向对比中,采用剪影图进行实验。为了便捷考虑,直接使用官方处理后的深度图像进行训练。点云数据具体处理过程如下:首先,点云数据包含每个点的三维坐标 (x, y, z) ,通过下述投影函数^[21]投影点云并离散化为二维点图:

$$r = \frac{a \tan 2(y, x)}{\Delta \theta} \tag{10}$$

$$c = \frac{\arcsin(z / \sqrt{x^2 + y^2 + Z^2})}{\Delta \varphi} \tag{11}$$

其中, $\Delta \theta$ 和 $\Delta \varphi$ 分别表示水平和垂直角分辨率。

每个二维图像的像素位置 (r, c) 对应于三维点的深度值 $d = \sqrt{x^2 + y^2}$ 。若多个点投影到一个二维位置,则保留距离最近的点,若没有三维点投影到特定的二维的位置,则填充 0。通过上述做法可以将点云数据从三维空间转换为二维平面的深度图像,再对深度图进一步做颜色映射。至此,完成了点云数据的处理工作,本文利用点云投影的

深度图进行训练,同时采用官方标准的数据划分方法,250 人作为训练集,其余样本作为测试集。

2.2 实验参数设置

本文的实验环境, GPU 为 NVIDIA GeForce RTX 3090, CPU 为 Inter® Core™ i9-12900k, 使用 Pytorch 深度学习框架。为了保证实验的可靠性,统一设定实验参数,考虑到硬件的负载能力,对于基准模型及相关实验,批次大小 (p, k, i) 分别设置为 $(8, 8, 10)$,其中, p 代表 ID 数量, k 代表每个 ID 步态序列数量, i 代表每个序列步态帧数量;SGD 作为优化器,权重衰减为 0.000 5;联合三元组损失和交叉熵损失监督模型训练,margin 设置为 0.2;初始学习率为 0.1,第 20 000 轮与 30 000 轮迭代时,学习率下降 0.1 倍,总迭代次数为 40 000 轮。

2.3 实验结果与分析

1)方法对比与实验分析

现阶段基于点云的步态识别研究较少,为了全面评估本文方法与其他先进方法的性能差异以及体现点云在步态识别中的优势,本文选取 LidarGait^[14] (CVPR2023) 和 HMRNet^[15] (ACM2024) 进行纵向比较。同时,选取多个基于二维剪影的先进步态识别方法进行横向比较,分别为 Gaitset^[3] (AAAI2019)、GaitPart^[4] (CVPR2020)、GaitBase^[1] (CVPR2023)。纵向比较输入的是点云,横向比较输入的是剪影。

由表 2 可知,基于三维点云的步态识别模型大幅优于基于二维剪影的步态识别模型,说明点云提供的高精度三维人体结构信息对步态识别有重要作用。同时,可以观察到,夜晚(NT)步态协变量在二维步态识别模型表现很差,但在点云步态识别中表现很好,其原因在于,摄像机受光线影响夜晚成像效果较差,而激光雷达采集点云却不受夜晚光线影响。

表 2 SUSTech1K 在不同模型的评估

Table 2 Evaluation of SUSTech1K in different models

模型	类型	Rank1 准确率								平均准确率	
		NM	BG	CL	CR	UB	UF	OC	NT	Rank1	Rank5
GaitSet ^[3]	剪影	70.93	70.23	38.55	66.84	66.48	63.89	68.68	24.38	67.01	85.78
GaitPart ^[4]		69.19	69.18	41.10	65.32	64.70	64.68	68.46	20.17	65.82	84.18
GaitBase ^[1]		80.05	77.83	49.20	75.49	75.64	77.54	81.25	25.21	75.87	89.19
HMRNet ^[15]		92.71	92.34	79.55	90.27	83.14	86.19	95.15	90.35	90.23	97.54
LidarGait ^[14]	点云	91.17	88.26	77.38	89.07	65.90	80.81	94.95	90.33	86.63	95.68
本文方法		95.40	92.93	75.08	93.16	88.09	91.80	96.49	90.59	92.65	98.19

在纵向比较上,本文方法除了衣服(CL)协变量外,其他所有步态协变量均显示出优越性,平均 Rank1 准确率达到 92.65%。与先进方法 LidarGait 相比,本文方法平均 Rank1 准确率整体提升 6.02%,特别是雨伞(UB)协变量,平均 Rank1 准确率从 65.90% 提升至 88.09%,提升了

22.19%,体现了本文方法的鲁棒性。

2)跨视角步态识别结果

为了准确评估模型跨视角步态识别性能,选取 LidarGait 跨视角步态识别结果与本文方法进行比较。本文实验遵循交叉视图评估协议^[14]。

从图 8 可以看出,本文方法在跨视角步态识别性能显著优于 LidarGait 方法。右侧混淆矩阵显示,本文方法在相同视角下的识别准确率普遍超过 96%,而 LidarGait 仅

在 92%左右。同时,本文方法在不同视角之间的识别准确率差距较小,显示出更强的跨视角鲁棒性。

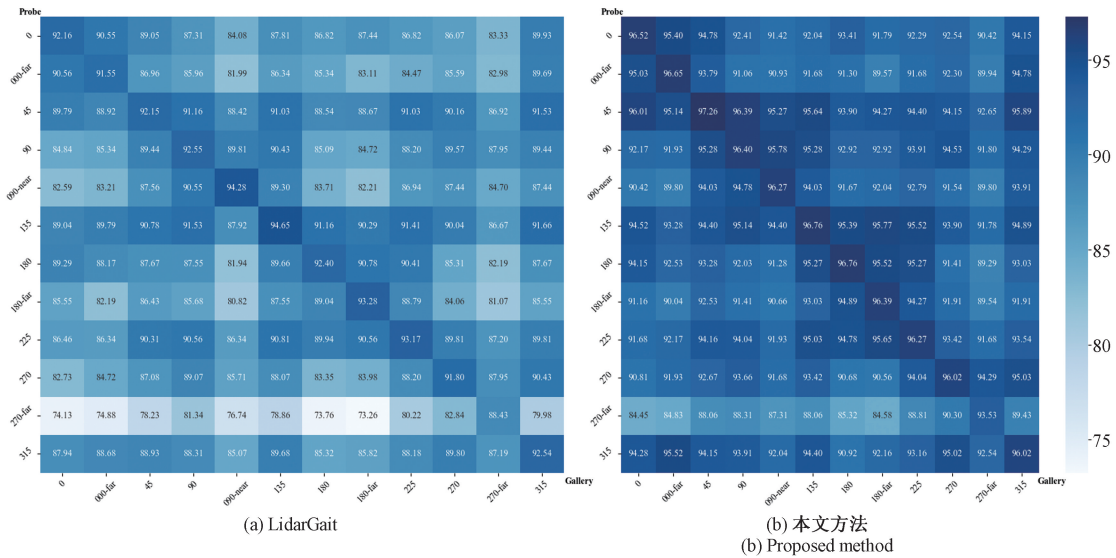


图 8 跨视角步态识别性能
Fig. 8 Cross-view gait recognition performance

值得注意的是,跨视角识别效果存在一定波动,相邻视角准确率较高,而随着视角差异增大,识别性能略有下降。可见,视角变化会显著增加模型学习一致性特征的难度。此外,视角‘270far’与其他视角匹配的准确率表现最差,主要原因是远距离垂直视角所包含的步态信息较少,不利于提取一致性特征。总的来看,本文方法在不同视角下的表现更加均衡,体现了本文方法在跨视角步态识别中的优越性。

2.4 消融实验

1)整体消融实验

本文提出的方法主要由三部分组成:首先,通过跨视角数据变换模块模拟不同视角的步态序列;其次,使用改进型残差网络提取全局多尺度的步态特征;最后,使用 KAN 网络增强局部细粒度步态特征的表征力。本文利用消融实验分析每个模块的实际作用。

从表 3 可以看出,使用改进型残差网络替换基础残差网络,平均 Rank1 准确率提升了 1.79%,说明改进型残差网络能提取高判别力的步态特征;紧接着在改进型残差网络的基础上添加跨视角数据变换模块,平均 Rank1 准确率提升了 3.69%,说明跨视角数据变换模块能较好提升模型的视角不变性和鲁棒性。同时,从表 4 可以观察到,在添加跨视角数据变换模块后,不同视角的识别准确率均有较大提升,进一步验证了其在适应视角变化方面的优越性。最后,加入 KAN 网络增强局部细粒度特征的表征力,平均 Rank1 准确率提升了 0.54%,说明非线性因素对局部细粒度特征再学习具有一定的作用。从图 9 也可以观察到,左侧每一分块的块内特征差异性不大,而右侧使用 KAN 网络对块内特征再学习,精简冗余特征,使特征表达更加紧凑,强化块内高判别力通道,使得步态特征更具代表性。

表 3 模型各模块有效性研究

模块	Table 3 Research on the effectiveness of each module of the model									%	
	Rank1 准确率									平均准确率	
	NM	BG	CL	CR	UB	UF	OC	NT		Rank1	Rank5
基础残差网络	91.17	88.26	77.38	89.07	65.90	80.81	94.95	90.33		86.63	95.68
+改进型残差网络	92.47	90.00	77.72	90.48	71.42	83.33	96.38	90.84		88.42	96.65
+跨视角数据变换模块	94.73	92.41	74.47	92.72	87.31	90.86	95.95	89.83		92.11	98.08
+KAN 局部细粒度增强网络	95.40	92.93	75.08	93.16	88.09	91.80	96.49	90.59		92.65	98.19

表 4 跨视角数据变换模块有效性研究

Table 4 Research on the effectiveness of Cross-view data transformation

模块	Rank1 准确率/%											
	0°	0°far	45°	90°	90°far	135°	180°	180°far	225°	270°	270°far	315°
基础残差网络	87.20	85.73	89.65	87.71	85.97	90.03	87.33	85.16	88.92	86.48	77.66	87.72
+改进型残差网络	88.74	88.61	90.86	89.03	87.88	90.92	88.70	87.45	90.13	87.67	81.54	89.54
+跨视角数据变换	92.14	91.47	94.68	93.35	91.93	94.09	92.28	90.54	93.04	92.47	86.40	92.93

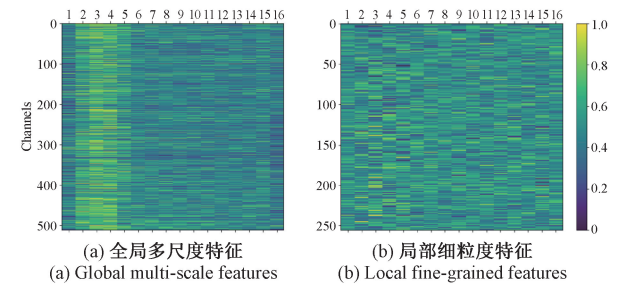


图 9 同一序列全局与局部特征可视化

Fig. 9 Visualization of global and local features of the same sequence

2)改进型残差网络消融实验

与基础残差网络相比,改进型残差网络主要是有 3 点改进。首先,混合使用 BasicBlock 和 Bottleneck 两个残差模块,其次,将 3×3 的初始卷积替换为多尺度的 Inception 模块,最后,统一将 ReLU 激活函数修改为 ReLU6。

从表 5 可以看出,添加混合残差网络平均 Rank1 准确率提高了 0.83%,继续添加 Inception 模块,Rank1 准确率进一步提高了 0.86%,最后将激活函数替换为 ReLU6,有 0.1%的提升。与基础残差网络相比,改进型残差网络在 Rank1 准确率上整体提升了 1.79%,说明该模块能捕捉更具判别力的步态特征。

表 5 改进型残差网络有效性研究

Table 5 Research on the effectiveness of each module of the model

残差网络	Rank1 准确率								平均准确率	
	NM	BG	CL	CR	UB	UF	OC	NT	Rank1	Rank5
基础残差网络	91.17	88.26	77.38	89.07	65.90	80.81	94.95	90.33	86.63	95.68
+混合残差网络	92.30	89.12	75.91	89.59	69.14	81.43	95.75	90.57	87.46	96.26
+Inception 模块	92.76	89.83	76.12	90.22	72.17	82.58	95.95	91.03	88.32	96.77
+ReLU6 激活函数	92.47	90.00	77.72	90.48	71.42	83.33	96.38	90.84	88.42	96.65

3 结 语

本文提出一种新的点云步态识别框架,通过跨视角数据变换模块模拟不同的步态序列,丰富样本的多样性,利用改进型残差网络提取全局多尺度步态特征,最后利用 KAN 网络增强局部细粒度步态特征的表征力。在 SUSTech1K 数据集的实验中充分验证了各个模块的有效性,与其他方法比较,本文提出的方法具有显著的优势。未来的研究可进一步考虑多模态融合策略,提升模型学习一致性特征的能力。

参考文献

[1] FAN CH, LIANG J H, SHEN CH F, et al. Opengait: Revisiting gait recognition towards better practicality[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition,2023: 9707-9716.

[2] 王茜,蔡竞,郭柏冬,等.面向公共安全的步态识别技术研究[J].中国人民公安大学学报(自然科学版),2023, 29(1):68-76.

WANG Q, CAI J, GUO B D, et al. Research on gait recognition technology for public security[J]. Journal of People' s Public Security University of China (Science and Technology),2023,29(1):68-76.

[3] CHAO H Q, WANG K, HE Y W, et al. GaitSet: Cross-view gait recognition through utilizing gait as a deep set[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2022, 44 (7): 3467-3478.

[4] FAN C, PENG Y J, CAO CH SH, et al. Gaitpart: Temporal part-based model for gait recognition[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2020.

[5] HUANG X H, ZHU D W, WANG H, et al. Context-sensitive temporal feature learning for gait recognition[C]. IEEE/CVF International Conference on Computer Vision, 2021:12909-12918.

[6] 邹雪,谭棉,严晓波,等.基于多尺度特征融合的跨视角

- 步态识别[J]. 电子测量技术, 2024, 47(1): 186-192.
- ZOU X, TAN M, YAN X B, et al. Cross-view gait recognition based on multi-scale feature fusion[J]. Electronic Measurement Technology, 2024, 47(1): 186-192.
- [7] 张红颖, 田鹏华. 结合残差网络与多级分块结构的步态识别方法[J]. 电子测量与仪器学报, 2022, 36(6): 66-72.
- ZHANG H Y, TIAN P H. Gait recognition method combining residual network and multi-level block structure[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(6): 66-72.
- [8] 李言, 曾维, 蒋毅, 等. 基于改进型对抗网络的步态特征提取方法研究[J]. 电子测量技术, 2022, 45(9): 121-126.
- LI Y, ZENG W, JIANG Y, et al. Research on gait feature extraction method based on improved generative adversarial networks[J]. Electronic Measurement Technology, 2022, 45(9): 121-126.
- [9] HUANG ZH, XUE D X, SHEN X, et al. 3D local convolutional neural networks for gait recognition[C]. IEEE/CVF International Conference on Computer Vision, 2021: 14920-14929.
- [10] TEEPE T, KHAN A, GILG J, et al. Gaitgraph: Graph convolutional network for skeleton-based gait recognition[C]. 2021 IEEE International Conference on Image Processing(ICIP), 2021: 2314-2318.
- [11] 沈澍, 张文昊, 丁浩, 等. 三维步态识别研究进展[J]. 中国图象图形学报, 2024, 29(7): 1921-1933.
- SHEN SH, ZHANG W H, DING H, et al. Research progress of three-dimensional gait recognition[J]. Journal of Image and Graphics, 2024, 29(7): 1921-1933.
- [12] FU Y, MENG SH B, HOU S H, et al. Gpgait: Generalized pose-based gait recognition[C]. IEEE/CVF International Conference on Computer Vision, 2023: 19595-19604.
- [13] ZHENG J K, LIU X CH, LIU W, et al. Gait recognition in the wild with dense 3d representations and a benchmark[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 20228-20237.
- [14] SHEN CH F, FAN CH, WU W, et al. Lidargait: Benchmarking 3d gait recognition with point clouds[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 1054-1063.
- [15] HAN X, REN Y M, CONG P SH, et al. Gait recognition in large-scale free environment via single LiDAR[C]. 32nd ACM International Conference on Multimedia, 2024: 380-389.
- [16] LIU Z M, WANG Y X, VAIDYA S, et al. KAN: Kolmogorov-arnold networks[J]. ArXiv preprint arXiv: 2404.19756, 2024.
- [17] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [18] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection[J]. ArXiv preprint arXiv: 2405.14458, 2024.
- [19] LIAO W B, ZHU Y H, WANG X Y, et al. Lightm-unet: Mamba assists in lightweight unet for medical image segmentation[J]. ArXiv preprint arXiv: 2403.05246, 2024.
- [20] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. ArXiv preprint arXiv: 1502.03167, 2015.
- [21] LI B, ZHANG T L, XIA T. Vehicle detection from 3D lidar using fully convolutional network[J]. ArXiv preprint arXiv: 1608.07916, 2016.

作者简介

魏永超, 博士, 教授, 主要研究方向为人工智能、计算机视觉和光电信息处理。

E-mail: weiyongchao@cafuc.edu.cn

谢卫鑫(通信作者), 硕士研究生, 主要研究方向为计算机视觉、模式识别、步态识别。

E-mail: xwx@cafuc.edu.cn

张娅岚, 副教授, 主要研究方向为人工智能、数字信号处理、计算机视觉。

E-mail: zyl@cafuc.edu.cn

王应海, 硕士研究生, 主要研究方向为计算机视觉、三维视觉缺陷检测。

E-mail: wyhelec@163.com

孙如新, 硕士研究生, 主要研究方向为计算机视觉、数字信号处理。

E-mail: srx@cafuc.edu.cn