

跨视角步态识别：一种融合跳跃连接与卷积块注意力模块的生成对抗网络方法

张红颖, 杨明川

(中国民航大学电子信息与自动化学院, 天津 300300)

摘要: 针对步态识别中因视角变化导致的识别率下降的问题, 本文提出一种融合跳跃连接与卷积块注意力模块的生成对抗网络(SC-GAN, generative adversarial network with skip connections and convolutional block attention module)的识别方法。首先, 在生成对抗网络中引入跳跃连接, 在通道域上进行特征图像的融合, 保留更多步态信息; 其次, 引入卷积块注意力模块, 以关注更多特定视角的特征位置和内容; 最后, 本文在中国科学院自动化研究所步态数据集B(CASIA-B, gait dataset B of the Institute of Automation, Chinese Academy of Sciences)上进行跨视角步态识别实验分析。实验结果表明, 本文模型在多重协变量因素影响下仍具有良好的鲁棒性和有效性。

关键词: 步态识别; 跨视角; 生成对抗网络(GAN); 跳跃连接; 卷积块注意力模块

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1674-5590(2026)02-0066-07

Cross-view gait recognition: a GAN approach with skip connections and convolutional block attention module

ZHANG Hongying, YANG Mingchuan

(College of Electronic Information and Automation, CAUC, Tianjin 300300, China)

Abstract: To address the problem of decreased recognition accuracy in gait recognition caused by viewpoint variations, this paper proposes a recognition method based on a generative adversarial network (GAN) with skip connections and a convolutional block attention module (SC-GAN). First, skip connections are introduced into the generative adversarial network to fuse feature maps in the channel domain, thereby preserving more gait-related information. Second, the convolutional block attention module is introduced to focus on more feature locations and content specific to particular viewpoints. Finally, cross-view gait recognition experiments are conducted on the gait dataset B of the Institute of Automation, Chinese Academy of Sciences (CASIA-B dataset). Experimental results demonstrate that the proposed model maintains good robustness and effectiveness under multiple covariate factors.

Key words: gait recognition; cross-view; generative adversarial network (GAN); skip connections; convolutional block attention module

步态识别凭借无需用户主动配合、非接触、无侵入性、难以伪装或隐藏, 及支持远距离识别等独特优势, 在实际应用场景中相较于指纹、人脸、虹膜等其他生物特征识别技术展现出更高的研究价值^[1]。

在步态识别研究中, 视角变化、携带物变化与衣着变化等协变量是影响识别性能的主要因素, 已受到国内外学者的广泛关注^[1]。其中, 视角变化所引起的跨视角问题, 是指因摄像机与行人相对角度变化而导致

识别率下降的现象, 该问题严重制约了步态识别在实际场景中的应用, 也因此成为领域内的研究重点。生成式方法^[2]将步态图像中的不同步态特征进行转化, 使之成为特定状态下的步态特征, 以实现跨视角步态识别。He 等^[3]提出了多任务生成对抗网络(MGAN, multi-task generative adversarial network), 在保留更多时序信息的同时, 有效降低了大视角变化下的累积转换误差。Hu 等^[4]提出了判别式步态生成对抗网络

(DiGGAN, discriminative gait generative adversarial network), 该模型可以将步态图像转换至不同视角, 并有效提取视角不变特征。Chen 等^[5]提出的多视角步态生成对抗网络 (MvGGAN, multiview gait generative adversarial network), 通过生成假步态样本以扩充数据集, 为跨视角步态识别提供充分数据支持。Yu 等^[6]提出了一种语义金字塔自编码器 (SPAEC, semantic pyramid autoencoder) 模型, 可将不同条件下的步态图像转换到统一特定状态下进行识别, 但其模型训练高度依赖标注数据。该团队后续提出的步态生成对抗网络^[7] (GaitGAN, gait generative adversarial network), 无需预先确定输入图像角度即可生成跨视角步态图像; 其改进的步态生成对抗网络第二版^[8] (GaitGANv2, gait generative adversarial network version 2) 进一步保留更多身份信息, 在不变特征提取方面实现了更优性能。

上述生成式方法虽然为跨视角步态识别研究提供了新思路并展现出优异性能。但其在携带物变化、衣着变化等多重协变量影响下的识别率仍会显著下降。考虑到实际应用中这些因素不可忽略, 如何提升多重协变量场景下的跨视角步态识别率, 仍是当前研究亟待解决的关键问题。

综上, 本文提出一种融合跳跃连接与卷积块注意力模块的生成对抗网络 (SC-GAN, generative adversarial network with skip connections and convolutional block attention module), 通过引入跳跃连接以改进网络结构, 同时结合卷积块注意力模块 (CBAM, convolutional block attention module) 抑制噪声干扰以优化生成的特征不变图像; 在中国科学院自动化研究所步态数据集 B (CASIA-B, gait dataset B of the Institute of Automation, Chinese Academy of Sciences) 上对 SC-GAN 进行跨视角步态识别实验, 并与多种先进方法进行对比分析, 结果验证了该方法的有效性与竞争力。

1 跨视角步态识别网络模型设计

本文的跨视角步态识别网络模型架构主要分为 3 个部分, 如图 1 所示。模型以多个视角下具有步态周期性特征的步态能量图 (GEI, gait energy image) 作为输入。在 SC-GAN 的结构设计中, 通过引入跳跃连接以保留更多步态信息, 并融合 CBAM 来关注更多特定视角的特征位置和-content。最后, 在特征降维和分类阶段, 选择主元分析 (PCA, principal component analysis) 和线性判别分析 (LDA, linear discriminant analysis) 相结合的策略。综合上述步骤, 构建不同视角下步态识

别的分类算法。



图 1 跨视角步态识别网络模型架构

Fig.1 Cross-view gait recognition network architecture

1.1 步态能量图

Hu 等^[9]提出的 GEI 具有周期性特点, 能够有效保留步态图像的时间信息, 并在单幅图像中描述人体在一个完整步态周期内的运动模式。得益于上述特点, 与其他步态图像表示方法相比, GEI 不仅节省了存储空间和计算时间, 还能通过周期加权平均的方式降低每帧轮廓中的噪声影响。

以 $B_t(x, y)$ 来表示在时间 t 上的步态图像序列, 则 GEI 定义为

$$G(x, y) = \frac{1}{N} \sum_{t=1}^N B_t(x, y) \quad (1)$$

式中: N 是一个完整步态周期的帧数; x 和 y 是 GEI 的坐标值。

每个 GEI, 既是人体步态的空间归一化能量表示, 又是该步态在一个完整步态周期内的时间归一化累积结果, 且其像素值通过灰度深浅直观反映步态运动的频率特性。因此, GEI 同时包含动态与静态信息, 可有效表征步态的时空特性。

本文将不同视角的正常行走 (nm)、携带物体行走 (bg)、穿着大衣行走 (cl) 的步态图像的 GEI 作为输入源数据。鉴于 90° 侧视图被验证为步态识别的最佳角度图^[9], 本文选择将 90° 正常行走 GEI (也称 90° 视角不变 GEI) 设为目标数据以期获得更高的识别率, 如图 2 所示。

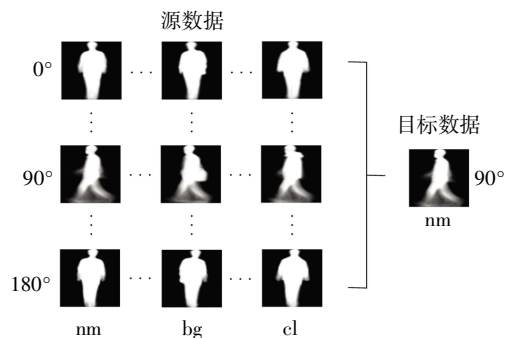


图 2 数据转换示意图

Fig.2 Data conversion diagram

1.2 SC-GAN

本节将按以下顺序介绍 SC-GAN 模型的构建过程: 首先介绍基础的 90° 视角不变 GEI 生成器; 接着给出与之配套的判别器结构; 随后为生成器引入跳跃连接; 最后进一步融合 CBAM, 形成完整的 SC-GAN 生

成器。需要说明的是,1.2.2节的判别器在整个训练过程中保持不变,用于与所有版本的生成器配对训练。由于各生成器输出的 GEI 图像尺寸和格式一致,该判别器可通用,从而确保模型性能的提升仅来自生成器结构的改进。

1.2.1 基础的 90°视角不变 GEI 生成器

现有实验数据直观表明,在 90°侧视图下的步态识别率最高^[4]。受像素级图像翻译模型 Pix2Pix(pixel-to-pixel,即“像素到像素”)^[9]的启发,本文提出将不同视角的步态图像统一映射至该最佳视角,以构建视角一致的步态表征。

本文所提出的 90°视角不变 GEI 生成器的基础模型如图 3 所示。该模型主要由对称的编码器和解码器两部分组成。编码器通过卷积(Convolution)层、批归一化(BatchNorm)层和 ReLU 激活函数依次级联构成经典下采样模块,逐步提取并压缩深层特征;解码器则通过反卷积(Deconvolution)层配合 BatchNorm 层、随机失活(Dropout)与 ReLU(或 Tanh)激活函数等操作实现特征上采样与空间重建,最终生成输出图像(输出数据)。

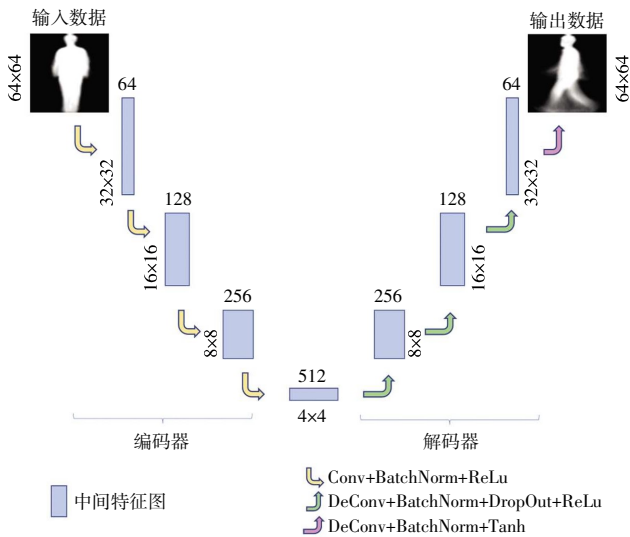


图 3 基础的 90°视角不变 GEI 生成器模型

Fig.3 Basic 90° view-invariant GEI generator model

1.2.2 90°视角不变 GEI 判别器

在生成对抗网络(GAN)中,判别器应尽可能准确地鉴别生成的“伪”GEI。在这一动态对抗博弈过程中,生成器被迫生成更加真实的 GEI,即保留更多特征信息,以利于后续特征提取与分类,提升识别准确率。为生成 90°视角不变 GEI,需对判别器进行训练以判断生成 GEI 的真伪。该判别器将输入图像(输入数据)映射为一个 0~1 之间的实数概率值,用以表征图像为

真的可能性。

本文所提出的 90°视角不变 GEI 判别器由多层卷积模块构成,网络前端采用 Convolution 层、BatchNorm 层与 ReLU 激活函数的经典结构进行级联,通过连续的卷积操作对输入特征进行下采样与空间压缩,特征图像尺寸减小的同时增加相应通道数。在网络的中间阶段,通过填充(Padding)层、Convolution 层、Batch-Norm 层与 ReLU 激活函数组成的模块进一步提取高层特征;在末端则通过 Padding 层、Convolution 层、BatchNorm 层与线性(Linear)层组成的分类头,将最终的特征图像映射为一个标量判别分数,用于评估输入数据属于真实样本的概率,从而为生成器的训练提供梯度反馈。GEI 判别器模型完整结构如图 4 所示。

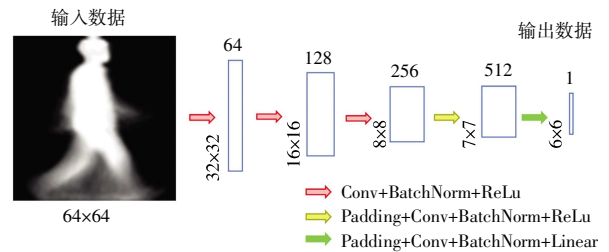


图 4 90°视角不变 GEI 判别器模型

Fig.4 90° view-invariant GEI discriminator model

1.2.3 引入跳跃连接的 90°视角不变 GEI 生成器

由图 3 可知,编码器最后一层输出结果为一张特征尺寸为 4×4 的特征图像,损失了很多特征信息,在一定程度上会影响最终生成图像的质量。此外,虽然编码器的浅层结构可以提取简单特征,但复杂特征则包含于深层结构之中。综上,本文借鉴 U-Net^[10]结构中的跳跃连接(skip connections)机制,将编码器与解码器对应位置的特征图像进行融合。该方法通过在通道维度上拼接特征,实现了深层高级语义特征与浅层低级细节信息的有效结合,从而形成信息更丰富、维度更高的复合特征图像。引入跳跃连接的 90°视角不变 GEI 生成器模型如图 5 所示。

1.2.4 融合跳跃连接和 CBAM 的 90°视角不变 GEI 生成器

CBAM 作为一种混合注意力机制,包括通道注意力模块和空间注意力模块两部分^[11],其结构如图 6 所示。CBAM 中通道注意力模块用于处理特征图像通道的分配关系,相当于一个特征检测器,聚焦有用特征信息的内容。而空间注意力模块用于关注 GEI 中对分类起关键作用的像素区域,忽略不相关的区域,聚焦有用特征信息的位置。因此,本文通过使用 CBAM 能够进一步提升模型的代表能力,保留更多有用特征

信息, 从而确保生成的 90° 视角不变 GEI 可以在分类识别中获得更高的准确率。融合跳跃连接和 CBAM 的 90° 视角不变 GEI 生成器模型如图 7 所示。

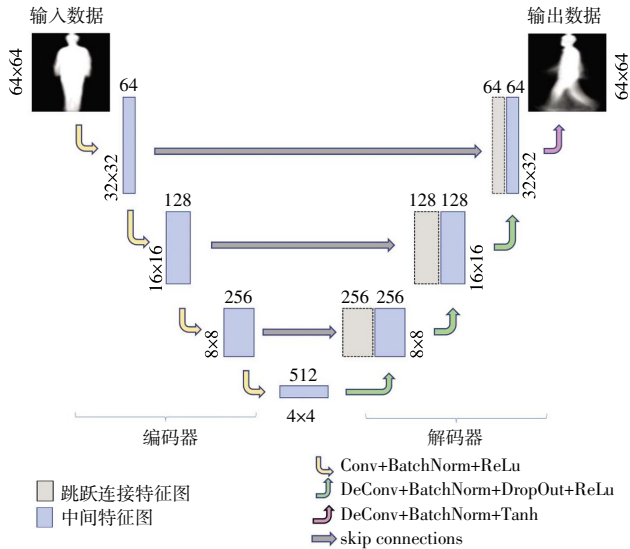


图 5 引入跳跃连接的 90° 视角不变 GEI 生成器模型

Fig.5 90° view-invariant GEI generator model with skip connections

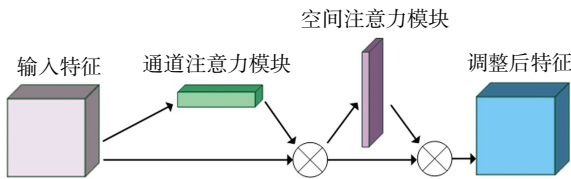


图 6 CBAM 结构

Fig.6 Structure of convolutional block attention module

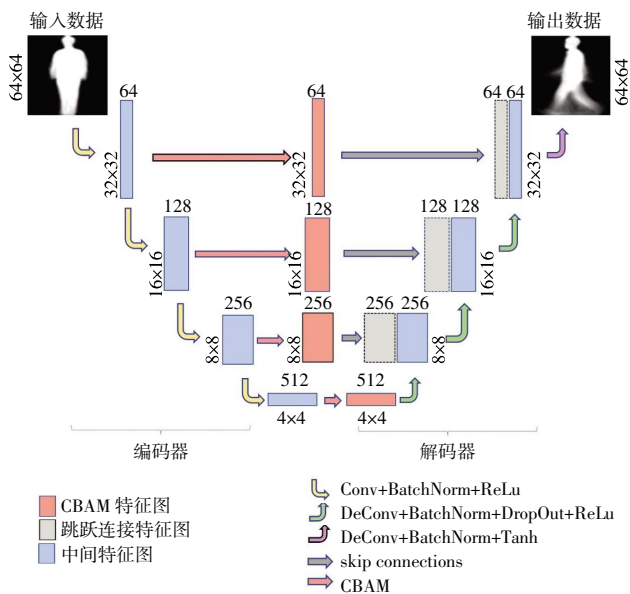


图 7 融合跳跃连接和 CBAM 的 90° 视角不变 GEI 生成器模型

Fig.7 90° view-invariant GEI generator model with skip connections and CBAM

1.3 特征降维和分类

PCA 是一种特征降维方法, 用于提取数据的主要

成分, 有效降低噪声与冗余; LDA 则可作为分类器, 利用特征信息优化分类性能。对于 GEI 而言, 其原始维度较高, 若直接用于分类, 易导致过拟合, 从而降低识别效果。因此, 通常先采用 PCA 进行降维, 在去除冗余特征的同时抑制噪声, 有效缓解过拟合问题; 随后将降维后的结果输入 LDA 进行判别, 以充分利用 LDA 最大化类间距离、最小化类内距离的特性, 从而提升模型的分类能力。

2 实验验证与分析

为充分验证 SC-GAN 的有效性, 在公认的步态识别基准数据集 CASIA-B 上进行跨视角步态识别实验。该数据集支持灵活的数据配置, 便于在公平条件下对先进方法进行性能对比。本实验的软硬件配置如下: 软件方面, 操作系统为 Windows, 深度学习框架为 TensorFlow 与 Keras, 编程语言为 Python。硬件方面, 使用 Intel Xeon(R) Silver 4112 CPU@2.6 GHz 处理器, 并搭载 2 张 NVIDIA GeForce RTX 2080 Ti 显卡。

2.1 实验设计

2.1.1 数据集描述

CASIA-B 数据集包含 124 名受试者, 每位受试者包含 11 个视角的步态图像, 视角范围覆盖 0°~180°, 相邻视角间隔为 18°。每个受试者提供 10 组行走序列, 包括 6 组正常行走 (nm-01 至 nm-06)、2 组携带物体行走 (bg-01, bg-02) 和 2 组穿着大衣行走 (cl-01, cl-02)。

图 8 展示了同一受试者在正常行走 (nm)、携带物体行走 (bg) 与穿着大衣行走 (cl) 3 种条件下, 共 11 个视角的步态图像序列。正常行走序列对应理想实验状态下的行人数据, 是模型性能验证的重要基准; 携带物体行走与穿着大衣行走 2 类序列则分别模拟携带物变化与衣着变化 2 种协变量因素, 用于检验模型对深层特征的学习能力。每位受试者的 11 个不同视角对应视角变化这一协变量, 是本文跨视角步态识别研究的核心因素。基于 CASIA-B 数据集生成的 GEI 如图 9 所示。

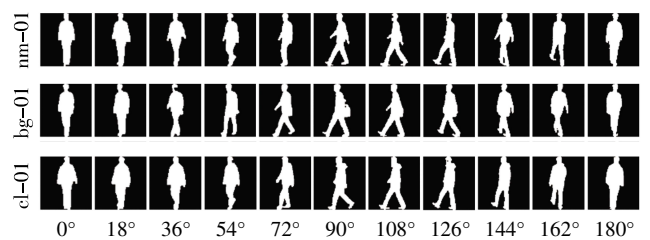


图 8 CASIA-B 数据集中不同视角下的步态图像序列示例

Fig.8 Examples of gait image sequences under different viewing angles in the CASIA-B dataset

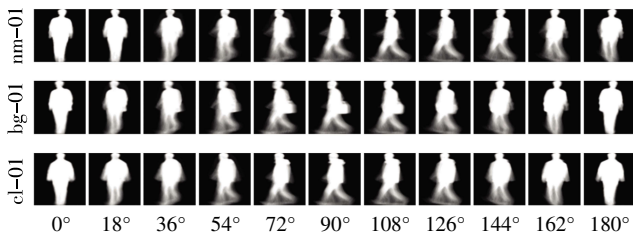


图9 基于CASIA-B数据集生成的GEI

Fig.9 GEI generated from the CASIA-B dataset

2.1.2 实验设置

本文参照图像识别领域中常用的数据集划分方法^[2],为步态识别任务定义了相应的训练集与测试集。其中,训练集用于模型的学习与参数优化,测试集则用于模型性能的评估,进一步地,测试集被划分为注册集和查询集^[2]。注册集由已知身份的步态图像构成,这些样本在测试前已完成身份标注,代表系统中已注册的个体;查询集则包含待识别的未知身份步态图像,其身份需通过与注册集中的样本进行比对来推断。

本文将CASIA-B数据集按受试者身份划分为互不重叠的训练集与测试集。

(1)训练集:由前62名受试者(ID 001-062)的所有可用步态图像序列构成,包括其6个正常行走序列(记为ID(001-062)-nm-01至ID(001-062)-nm-06)、2个携带物体行走序列(ID(001-062)-bg-01, ID(001-062)-bg-02)及2个穿着大衣行走序列(ID(001-062)-cl-01, ID(001-062)-cl-02)。这些数据仅用于模型训练。

(2)测试集:由后62名受试者(ID 063-124)的数据组成,并进一步划分为注册集与查询集:①注册集,使用每位受试者的前4个正常行走序列(记为ID(063-124)-nm-01至ID(063-124)-nm-04);②查询集,包含其余序列,即后2个正常行走序列(ID(063-124)-nm-05, ID(063-124)-nm-06)、2个携带物体行走序列(ID(063-124)-bg-01, ID(063-124)-bg-02)和2个穿

着大衣行走序列(ID(063-124)-cl-01, ID(063-124)-cl-02)。

2.1.3 模型结构与参数

1)生成器结构

生成器采用编码器-解码器结构。在编码阶段,通过卷积操作将特征图像的采样率逐层减半;相应地,在解码阶段则采用相同的上采样因子逐步恢复图像分辨率。本文所设计的生成器结构如下:

- 编码器 C64-C128-C256-C512
- 解码器 C512-C256-C128

其中,C表示卷积层,其后数值为相应通道数。

2)判别器结构

经过训练的判别器将用来判断生成器输出中某个区域是否为真。本文所设计的判别器结构如下:

- 判别器 C64-C128-C256-C1

2.2 基于CASIA-B的实验结果

本文SC-GAN的性能由正确分类率(CCR, correct classification rate)衡量,其定义为被正确分类的样本数量与总样本数量之比。

表1、表2和表3展示了SC-GAN在CASIA-B数据集上的步态识别性能。实验结果表明:①当查询集为正常行走序列时(表1),CCR在同视角下最高,并随视角差异增大而明显下降;②当查询集分别为携带物体行走序列(表2)或穿着大衣行走序列(表3)时,受外观变化干扰,整体CCR显著降低,且无法再观察到视角差异越大、CCR越低的清晰趋势。

2.3 消融实验

在CASIA-B数据集上,本文以GAN为基础框架,设计了3组消融实验:第一组为原始GAN基线;第二组在此基础上仅引入跳跃连接;第三组则进一步融合跳跃连接与CBAM。为全面评估SC-GAN在跨视角步态识别任务中的性能,实验充分利用了CASIA-B中涵

表1 查询集为正常行走序列的CCR

Tab.1 CCR for the probe set of normal-walking sequences

测试集	查询集(ID(063-124)-nm-05, ID(063-124)-nm-06)视角											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集(为ID(063-124)-nm-01至ID(063-124)-nm-04)视角	0°	97.58	66.13	50.81	44.35	29.84	23.39	24.19	28.23	38.71	47.58	75.00
	18°	77.42	99.19	92.74	71.77	49.19	38.71	42.74	47.58	52.42	54.84	52.42
	36°	50.00	91.13	98.39	93.55	73.39	59.68	60.48	67.74	68.55	54.84	35.48
	54°	37.90	69.35	92.74	96.77	93.55	86.29	85.48	80.65	73.39	47.58	27.42
	72°	25.00	39.52	72.58	91.94	99.19	97.58	96.77	86.29	66.13	38.71	21.77
	90°	22.58	30.65	54.03	79.84	97.58	99.19	99.19	83.87	54.03	34.68	20.16
	108°	25.00	41.13	63.71	83.87	95.16	96.77	99.19	94.35	76.61	49.19	23.39
	126°	20.16	52.42	60.48	76.61	83.06	81.45	92.74	97.58	96.77	55.65	27.42
	144°	35.48	55.65	66.13	68.55	68.55	59.68	77.42	97.58	99.19	80.65	38.71
	162°	53.23	66.13	61.29	49.19	36.29	39.52	41.13	60.48	88.71	99.19	75.00
	180°	70.97	45.97	31.45	20.97	19.35	18.55	17.74	24.19	33.87	75.00	99.19

%

表 2 查询集为携带物体行走序列的 CCR
Tab.2 CCR for the probe set of carrying-walking sequences

测试集	查询集(ID(063-124)-bg-01, ID(063-124)-bg-02)视角											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集(为 ID (063-124)-nm-01 至 ID(063-124)-nm-04) 视角	0°	84.68	54.03	36.29	25.00	18.55	21.77	17.74	19.35	25.00	33.87	51.61
	18°	58.06	88.71	75.00	48.39	39.52	31.45	33.06	38.71	45.97	44.35	36.29
	36°	38.71	79.03	84.68	70.97	55.65	40.32	37.10	43.55	52.32	42.74	25.81
	54°	29.84	55.65	81.45	86.29	79.64	59.68	57.26	55.65	57.26	38.71	25.00
	72°	21.77	33.87	54.84	75.00	93.55	82.26	79.64	70.16	45.16	27.42	21.77
	90°	18.55	33.87	50.00	60.48	83.06	84.68	81.45	63.71	36.29	25.81	21.77
	108°	15.32	33.06	49.19	60.48	79.64	79.03	87.90	79.64	68.55	30.65	15.32
	126°	24.19	31.45	44.35	55.65	63.71	58.87	75.00	82.26	77.42	41.13	23.39
	144°	29.84	40.32	47.58	41.94	49.19	47.58	54.03	76.61	84.68	55.65	35.48
	162°	37.90	40.32	36.29	29.03	27.42	30.65	27.42	39.52	58.06	82.26	50.61
	180°	52.23	29.03	25.81	20.97	12.90	11.29	9.68	13.71	27.42	41.13	87.90

表 3 查询集为穿着大衣行走序列的 CCR
Tab.3 CCR for the probe set of coat-walking sequences

测试集	查询集ID(063-124)-cl-01, ID(063-124)-cl-02视角											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集(为 ID (063-124)-nm-01 至 ID(063-124)-nm-04) 视角	0°	41.94	26.61	19.35	16.94	12.90	12.10	9.68	16.13	19.35	21.77	25.81
	18°	25.81	45.97	47.58	28.23	20.97	20.16	17.74	19.35	26.61	22.58	18.55
	36°	24.19	41.94	58.87	53.23	29.03	20.97	20.16	32.26	29.84	28.23	17.74
	54°	18.55	29.84	49.19	58.87	47.58	38.71	43.55	37.10	32.26	17.74	12.10
	72°	22.58	27.42	36.29	47.58	63.71	51.61	48.39	48.39	33.06	16.94	8.87
	90°	19.35	25.81	31.45	41.94	60.48	55.65	50.81	47.58	32.26	20.97	10.48
	108°	18.55	29.03	35.48	45.97	59.68	50.00	60.48	58.87	40.32	25.81	8.06
	126°	10.48	25.00	30.65	34.68	46.77	41.13	45.97	60.48	55.65	30.65	12.90
	144°	14.52	24.19	29.03	37.10	30.65	25.00	37.90	53.23	58.06	40.32	16.13
	162°	16.13	20.16	29.84	20.97	19.35	12.10	16.94	28.23	36.29	42.74	25.00
	180°	15.32	11.29	9.68	8.87	5.65	7.26	7.26	12.90	16.13	27.42	41.13

盖的所有视角的步态图像,并以 nm、bg、cl 3 种行走条件下 CCR 的平均值作为主要评价指标。消融实验结果如表 4 所示。

表 4 3 种行走条件下各消融变体的平均 CCR

Tab.4 Average CCR of ablation variants under three walking conditions

方案	CCR/%		
	nm	bg	cl
GAN	55.36	36.19	19.72
GAN+跳跃连接	61.03	45.27	28.39
GAN+跳跃连接+CBAM	62.22	48.10	30.44

2.3.1 引入跳跃连接的方法

由表 4 可知,在 GAN 上引入跳跃连接之后,3 种行走条件下的平均 CCR 均得到提升。引入跳跃连接之后的 GAN 通过通道域上的特征图像融合,有效缓解了解码器在上采样过程中因信息流失导致的特征重建不足问题。该结构实现了跨层级的特征复用,使解码器在上采样时能够融合来自编码器多层次、多尺度的细节信息,从而显著增强了生成图像的细节保留

能力。特别是存在协变量干扰的场景下,对细节信息捕获的需求更为突出,而跳跃连接机制恰好为此提供了有效保障。

2.3.2 融合跳跃连接和 CBAM 的方法

在引入跳跃连接的基础上,本文 SC-GAN 进一步引入 CBAM 以提升步态识别性能。实验结果表明,融合跳跃连接与 CBAM 的方案,在 3 种行走条件下的平均 CCR 均高于基线 GAN 及仅引入跳跃连接的 GAN。进一步引入的 CBAM 使 SC-GAN 在特征重建过程中同时兼顾重要通道与关键区域,增强了细节保留与判别能力。实验结果表明,融合跳跃连接与 CBAM 的 SC-GAN 有效提升了跨视角步态识别的性能。

2.4 对比实验

为全面评估 SC-GAN 的性能,本文在相同实验环境下,将 SC-GAN 与当前跨视角步态识别领域的先进方法进行了对比分析。参照相关文献的典型实验配置,选取了 SPAE^[6]、GaitGANv1^[7]与 GaitGANv2^[8]作为对比方法,在 nm、bg、cl 3 种行走条件下,开展了步态识

别性能的比较实验。

图 10 为注册集与查询集在同视角的设置下,不同方法在 nm、bg、cl 3 种行走条件下的平均 CCR。

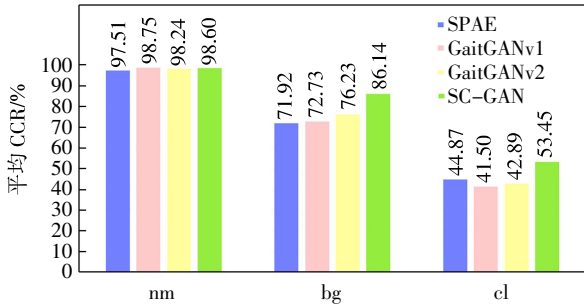


图 10 注册集和查询集在同视角下的平均 CCR

Fig.10 Average CCR under same-view setting between gallery and probe sets

本文进一步对比了注册集与查询集在跨视角的设置下,各方法在 nm、bg、cl 3 种行走条件下的平均 CCR,结果如图 11 所示。

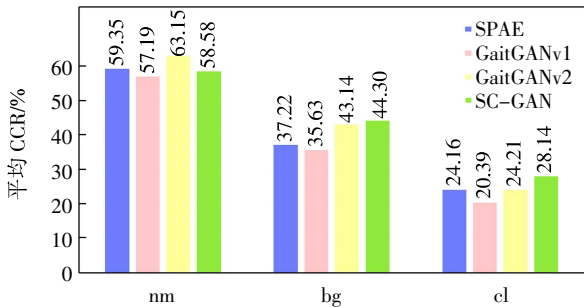


图 11 注册集和查询集在跨视角下的平均 CCR

Fig.11 Average CCR under cross-view setting between gallery and probe sets

根据图 10 与图 11,在正常行走条件下,本文的 SC-GAN 的平均 CCR 与现有先进方法表现相当;在携带物体行走和穿着大衣行走 2 种协变量条件下,SC-GAN 则实现了更高的平均 CCR。该结果表明,在多重协变量并存的跨视角场景中,SC-GAN 通过跳跃连接与 CBAM 实现了判别性细节特征的保留与生成能力的增强,有效提升了模型的识别鲁棒性,进一步印证了 SC-GAN 的有效性。

3 结语

本文提出了一种融合跳跃连接与 CBAM 的 SC-GAN 方法,旨在解决视角变化及携带物变化、衣着变化等协变量干扰下的步态识别难题。通过跳跃连接实现跨层级特征复用,有效保留细节信息;结合 CBAM 在通道与空间维度强化关键特征,提升生成的 90°视

角不变 GEI 的质量。在 CASIA-B 数据集上的消融实验表明,所提方法在正常行走(nm)、携带物体行走(bg)和穿着大衣行走(cl) 3 种条件下均显著优于基线 GAN。进一步与 SPAE、GaitGANv1、GaitGANv2 等先进方法的对比实验表明,SC-GAN 在同视角与跨视角场景下均具有更强的鲁棒性,尤其在复杂协变量条件下识别性能提升明显,验证了其有效性与竞争力。

参考文献:

- [1] 王科俊,丁欣楠,邢向磊,等.多视角步态识别综述[J].自动化学报,2019,45(5):841-852.
- [2] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [3] HE Y W, ZHANG J P, SHAN H M, et al. Multi-task GANs for view-specific feature learning in gait recognition[J]. IEEE Transactions on Information Forensics and Security, 2019, 14(1): 102-113.
- [4] HU B Z, GUAN Y, GAO Y, et al. Robust cross-view gait recognition with evidence: a discriminant gait GAN (DiGGAN) approach[EB/OL]. (2018-11-26)[2024-10-10]. <https://arxiv.org/abs/1811.10493>.
- [5] CHEN X, LUO X Z, WENG J, et al. Multi-view gait image generation for cross-view gait recognition[J]. IEEE Transactions on Image Processing, 2021, 30: 3041-3055.
- [6] YU S Q, CHEN H F, WANG Q, et al. Invariant feature extraction for gait recognition using only one uniform model[J]. Neurocomputing, 2017, 239: 81-93.
- [7] YU S Q, CHEN H F, REYES E B G, et al. GaitGAN: invariant gait feature extraction using generative adversarial networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 532-539.
- [8] YU S Q, LIAO R J, AN W Z, et al. GaitGANv2: invariant gait feature extraction using generative adversarial networks[J]. Pattern Recognition, 2019, 87: 179-189.
- [9] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 5967-5976.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention: MICCAI 2015. Cham: Springer, 2015: 234-241.
- [11] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision: ECCV 2018. Cham: Springer, 2018: 3-19.
- [12] GAO W, CAO B, SHAN S G, et al. The CAS-PEAL large-scale Chinese face database and baseline evaluations[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2008, 38(1): 149-161.

(责任编辑:刘智勇)