

融合风格编码的双通道生成对抗网络步态识别

周 航¹, 宋珂蔚¹, 徐源昊², 陈业泓¹, 姜 江¹

(1. 北京交通大学 电子信息工程学院, 北京 100044; 2. 深圳市路桥建设集团有限公司 市政工程分公司, 广东深圳 518024)

摘 要: 针对现有步态识别技术存在的视角适应性差、模型训练稳定性不足的问题, 提出了一种融合风格编码的双通道时空卷积生成对抗网络模型。该模型通过双通道时空卷积网络分别提取步态图像中的复杂时空特征与步态特征, 并引入风格编码器来减少生成图像与原图之间的语义差异, 从而提升生成器的稳定性。基于 CASIA-B 数据集的实验结果表明, 本文模型相比 GaitGAN 在 54°、90° 和 126° 这 3 个具有代表性的步态视角下, 识别率分别提高了 23%、28% 和 22%。此外, 该方法在 OU-MVLP 大规模跨视角步态数据库中仍展现出较好的适用性, 达到了 89.1% 的平均识别精度, 有效提高了跨视角步态识别率。

关键词: 步态识别; 跨视角; 时空卷积网络; 风格编码器; 生成对抗网络

中图分类号: TP391 **文献标志码:** A **文章编号:** 1671-5497(2026)07-1970-14

DOI: 10.13229/j.cnki.jdxbgxb.20250002

Gait recognition using dual-channel generative adversarial network with style encoding integration

ZHOU Hang¹, SONG Ke-wei¹, XU Yuan-hao², CHEN Ye-hong¹, JIANG Jiang¹

(1. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China; 2. Municipal Engineering Branch, Shenzhen Luqiao Construction Group Co., Ltd., Shenzhen 518024, China)

Abstract: Aiming at the problems of poor cross-view adaptability and insufficient model training stability existing in current gait recognition technologies, a dual-channel spatiotemporal convolutional generative adversarial network (GAN) model that incorporates style encoding was proposed. The model employs a dual-channel spatiotemporal convolutional network to extract complex spatiotemporal and gait features from gait images, while a style encoder is introduced to minimize the semantic discrepancies between the generated images and the original ones, thereby enhancing the stability of the generator. Experimental results on the CASIA-B dataset demonstrate that, compared to GaitGAN, the proposed model achieves recognition rate improvements of 23%, 28%, and 22% at three representative gait angles of 54°, 90°, and 126°, respectively. Additionally, the method shows strong applicability on the OU-MVLP large-scale cross-view gait database, achieving an average recognition accuracy of 89.1%, significantly improving cross-view gait recognition performance.

收稿日期: 2025-01-02.

基金项目: 国家自然科学基金项目(61872027); 国铁集团重点项目(M23D00101); 北京交通大学科研项目(W21L00390); 北京交通大学教改项目(SQ20240024).

作者简介: 周航(1974-), 男, 副教授, 博士. 研究方向: 数字图像处理. E-mail: hangzhou@bjtu.edu.cn

Key words: gait recognition; cross-view; spatio-temporal convolutional network; style encoder; generative adversarial network

0 引言

近年来,步态识别作为新兴生物识别技术,因其无需主动配合、可非接触式获取特征等特点,在人机交互、智能安防和医疗诊断等领域展现出广泛的应用前景^[1-4]。然而,视角变化是影响识别性能的关键因素,直接影响步态特征的全局表现;衣着变化和携带物品等因素也增加了识别难度,导致跨视角识别准确性下降^[5]。

随着深度学习的发展,学者们开始采用卷积神经网络^[6]、循环神经网络^[7]、自编码器^[8]等方法进行跨视角步态识别。然而,基于有监督学习的方法依赖大量人工标注,成本高且耗时,限制了模型的泛化能力。因此,开发无需大量标注数据的无监督学习模型成为重要课题。无监督学习通过分析数据内在特征,完成识别任务,如基于生成对抗网络(Generative adversarial network, GAN)^[9]、聚类^[10]和无监督适应^[11]的步态识别方法。基于GAN的步态识别因其生成高质量样本和自动学习特征,已成为研究热点。

尽管基于GAN的步态识别模型带来了创新的改进,然而GaitGAN^[12]在图像质量和时序建模上存在不足,生成的步态图像缺乏细节和清晰度,且时序一致性较差,无法有效捕捉步态的动态特征。为了提高跨视角步态识别的准确率,解决实验数据不足、生成器在训练过程中易崩溃等问题,本文提出了一种融合风格编码的双通道时空卷积生成对抗网络模型(Dual-channel spatio-temporal convolutional generative adversarial network with style encoding, DSCGAN-SE)。该模型的生成器采用双通道时空卷积网络设计,利用多尺度一维卷积核分别提取步态图像的横向时序特征和纵向空间特征,并通过跨尺度的局部特征融合增强特征表达。最终,时空融合特征被引入残差模块进行全局特征提取,从而更加有效地生成既保留步态特征又保持原始图像语义信息的图像。此外,为了更好地在跨视角步态图像生成中保持身份信息,DSCGAN-SE引入了风格编码器,使生成器生成的图像信息最大化接近真实图像。实验结果表明,基于CASIA-B^[13]数据集,本文模型

在54°、90°和126°三个具有代表性的步态视角下的识别率分别较基准模型GaitGAN提升了23%、28%和22%。同时,该模型在OU-MVLP^[14]大规模跨视角步态数据库上展现了良好的适用性,平均识别精度达到89.1%。实验结果充分证明了该模型在应对大规模数据集复杂视角变化时的优越性能,为实际应用提供了可靠的技术支持。

1 基于GAN的步态识别模型

GAN是一种由Goodfellow等^[9]于2014年提出的深度学习架构,旨在通过对抗性训练机制生成高质量的数据样本。GAN主要由两个核心部分组成:生成器(Generator, G)和判别器(Discriminator, D)。生成器的功能是生成具有特定特征的样本数据,以便“欺骗”判别器,而判别器的功能则是区分输入的数据是来自生成器还是实际的真实数据。

1.1 GaitGAN的网络架构

GaitGAN^[12]是一种基于GAN的步态识别模型,旨在通过生成器-判别器架构生成逼真的跨视角步态图像。生成器通过学习步态特征,生成与目标视角匹配的步态图像,而判别器则用于区分生成的步态图像与真实图像。GaitGAN旨在生成标准化视角图像(通常为90°视角),以此作为匹配和识别的基础,从而开创了利用GAN技术进行步态识别的新篇章,具体研究思路如图1所示。

与传统生成对抗模型的结构不同,GaitGAN的网络架构主要由生成器G和2个具备不同功能

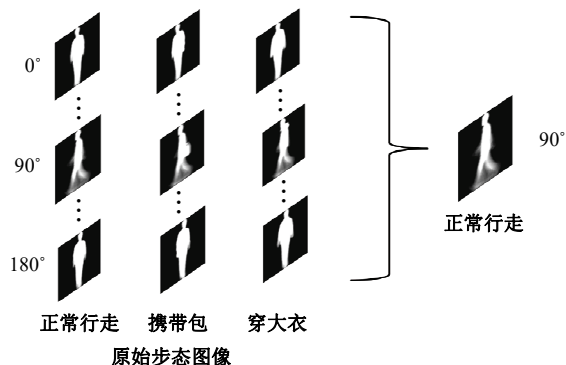


图1 GaitGAN原理示意图

Fig. 1 Schematic diagram of the GaitGAN principle

的判别器(真假判别器 D_R 和领域判别器 D_A)组成,如图2所示。

GaitGAN的生成器 G 由编码器和解码器组成,其中编码器通过4个卷积层结合批量归一化和LeakyReLU激活函数,负责从源图像中提取特征;解码器则利用4个分数步长卷积层将特征映射重建为图像。

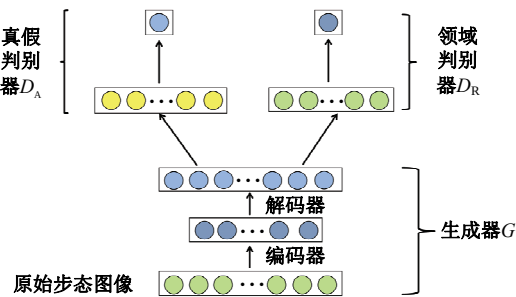


图2 GaitGAN的网络结构示意图

Fig. 2 Schematic diagram of the GaitGAN network structure

真假判别器 D_R 用于评估图像的真实性,网络结构包括卷积层、池化层和批量归一化层,通过ReLU激活函数和最大池化减少特征图的空间尺寸,并通过全连接层输出真实性概率。领域判别器 D_A 则专注于身份识别,评估生成的姿态序列与已知身份特征的相似度,最终通过全连接层输出身份标签或特征度量值,以确保模型能够准确识别个体身份。

尽管GaitGAN在步态生成领域取得了一定成果,但模型设计上仍存在局限性。首先,生成器主要依赖于生成的侧视图进行特征匹配,未能充分提取中间层及其他潜在有效特征,同时忽视了步态的时序特征。其次,GaitGAN采用的单一损失函数未能整合多样的判别标准,在生成图像与原始步态图像的语义差异方面表现不足,影响了模型的判别准确性。为此,DSCGAN-SE模型通过优化生成器,更全面地捕捉步态及其时空特征,并设计了更复杂的损失函数,以更有效反馈生成图像与真实图像之间的语义差异,从而提升了模型的性能。

1.2 DSCGAN-SE网络架构

尽管GaitGAN在步态图像生成任务中展现了良好的性能,但生成器在图像质量和生成过程的稳定性方面仍面临挑战。具体而言,生成器常出现图像质量较低的问题,这限制了GaitGAN模型在实际应用中的效果和可靠性。为克服此类问

题,本文在GaitGAN的基础上对生成器进行优化,提出DSCGAN-SE模型以提升图像质量和生成稳定性,并融入了风格编码器,改善了生成图像与真实图像之间的差异。

所设计的DSCGAN-SE模型总体结构如图3所示。在该模型中,CASIA-B^[13]数据集经过预处理后被划分为3个数据集类别:相关图集 $I_{\text{ass-label}}$ 、非相关图集 $I_{\text{noass-label}}$ 和源图集 I_{img} 。 $I_{\text{ass-label}}$ 与 I_{img} 用于识别不同个体在相同条件和角度下的动作差异; $I_{\text{noass-label}}$ 与 I_{img} 则用于识别同一人在不同条件和角度下的动作变化,目的是评估模型在不同动作条件下对同一人的识别效果。

通过这种分类方法,可以确保生成器在训练过程中从步态数据集中采样到具有不同身份ID、状态和角度的步态图像,从而提高模型对多样化步态特征的生成能力。

DSCGAN-SE整体生成框架由生成器 G 、真实判别器 D_{op} 、身份判别器 D_{ip} 、风格编码器构成。在DSCGAN-SE模型中,本文首次提出风格编码器,用于从源图集 I_{img} 的步态图像和生成器生成的步态假图像 I_{fake} 中提取风格特征,并通过3个关键步骤实现不同视角和状态下的步态图像生成。具体过程如下:

(1)生成器 G 从源图集 I_{img} 随机选择一张步态图像作为双通道时空卷积网络的输入,以提取图像中的步态和时空特征,目的是生成步态假图像 I_{fake} 。随后,将步态假图像 I_{fake} 和真实图像 I_{real} 输入风格编码器中提取两者的特征信息,并计算其差异值进行损失回传,用于训练过程的迭代更新,公式表达为:

$$I_{\text{fake}} = G(E(I_{\text{img}})) \quad (1)$$

(2)真实判别器 D_{op} 将生成器 G 生成的步态假图像 I_{fake} 与相关图集 $I_{\text{ass-label}}$ 和非相关图集 $I_{\text{noass-label}}$ 作为输入,分别对假图像和两张真图像进行判别,对判别结果计算交叉熵损失,将损失分别回传用来迭代更新,以不断提升区分真伪图像的能力。

在身份判别器 D_{ip} 中,首先将源图集 I_{img} 分别与生成器 G 生成的步态假图像 I_{fake} 、相关图集 $I_{\text{ass-label}}$ 、非相关 $I_{\text{noass-label}}$ 进行融合,将融合后的新集合作为输入,分别对是否与源图集 I_{img} 具有相关性进行真假判别,对判别结果计算交叉熵损失进行误差回传,以确保生成器生成的图像尽可能保留原始图像的语义信息。

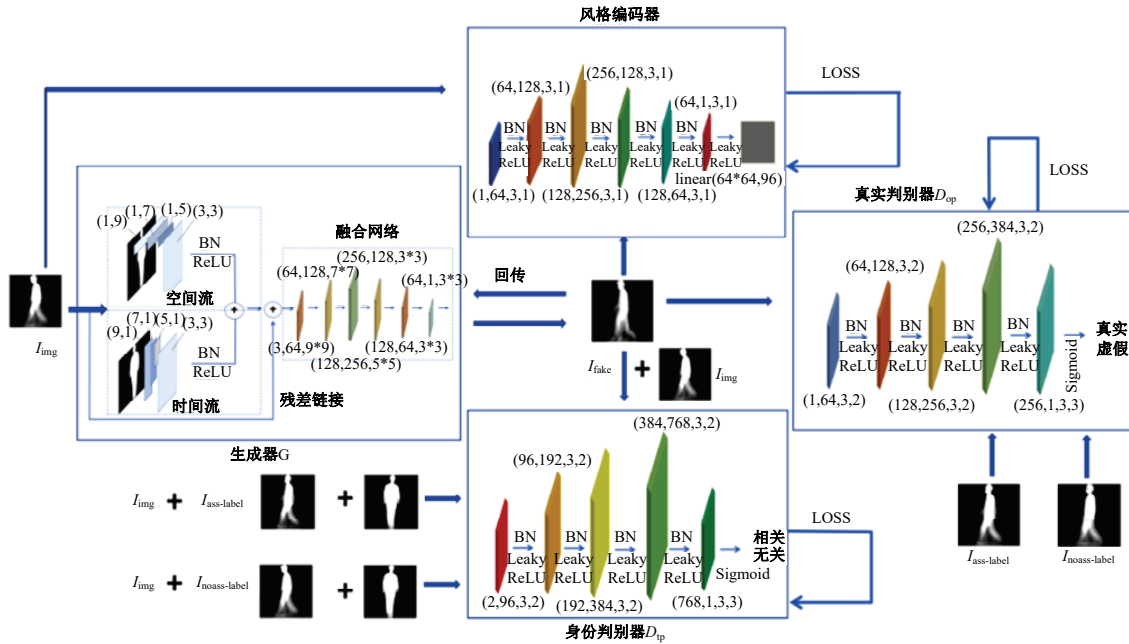


图 3 DSCGAN-SE 的总体网络框架

Fig. 3 Overall network framework of the DSCGAN-SE

(3)随着训练迭代的进行,对抗训练逐渐趋向平衡,即真实判别器 D_{op} 再无法有效区分生成器生成的图像与真实图像。在测试阶段,将原相关图集 $I_{ass-label}$ 、非相关图集 $I_{noass-label}$ 和源图集 I_{img} 同时输入训练好的生成器 G ,以生成最终的步态图像,达到扩展数据集的目的,用公式可表达为:

$$I_{fake} = G(E(I_{ass-label}, I_{noass-label}, I_{img})) \quad (2)$$

DSCGAN-SE 模型首次引入风格编码器,从源步态图像和生成的步态假图像中提取风格特征,通过计算两者之间的特征差异优化训练过程。根据生成器生成跨视角步态图像,并结合对抗性训练机制,DSCGAN-SE 能够有效提升跨视角步态识别的精度,同时实现数据集的扩展。

2 本文模型

2.1 生成器网络架构

为生成高质量的步态图像,同时增强生成器 G 在训练过程中稳定性,DSCGAN-SE 在双流卷积网络^[19]的基础上,设计了双通道时空卷积网络,将其融合至生成器 G 的网络结构设计当中,主要针对源图集 I_{img} 中步态和时空特征提取,同时引入残差融合网络,逐步恢复步态假图像 I_{fake} 的空间细节和结构信息。其网络结构如图 4 所示,生成器 G 的网络结构参数设置如表 1 所示。

在双通道时空卷积网络中,时间流卷积网络对源图集 I_{img} 中步态图像进行步态特征提取,采用

了 3 种不同尺寸的横向一维卷积核 (1×9 、 1×7 、 1×5),用于捕捉步态图像中的多尺度信息。随后,这些一维卷积层通过 ReLU 激活函数进行非线性变换,对各层输出拼接后经过一个 3×3 的二维卷积层,以整合不同卷积尺度的步态特征。

空间流卷积网络专注于源图集 I_{img} 中步态图像中空间特征的提取,该网络设计包含三层纵向一维卷积核 (9×1 、 7×1 、 5×1),每个卷积层的输出经过 ReLU 激活函数后进行通道级拼接,以融合多尺度的时序信息。拼接后的特征图通过一个 3×3 的二维卷积层进行综合处理,以整合和优化提取的时序特征。最后,通过 ReLU 激活函数和一个 3×3 的二维卷积层提取空间特征。最终,通过批归一化 (Batch normalization, BN) 和 ReLU 激活函数对特征图进行标准化和非线性变换,从而提升提取的时序特征稳定性。

残差融合网络采用转置卷积操作整合空间流和时间流的图像特征,旨在将空间流和时间流网络中提取的步态和时序特征映射至高分辨率输出,通过一系列反卷积层逐步上采样并重建图像。残差融合网络包含 6 个反卷积层,每层均配备批归一化 BN 和 ReLU 激活函数,用于逐步增加生成步态假图像 I_{fake} 的分辨率。反卷积层依次将特征图的通道数扩大,并通过逐步采样至目标尺寸 64×64 ,最终生成符合目标尺寸的步态图像。

转置卷积的优势在于其能够逐步恢复生成步

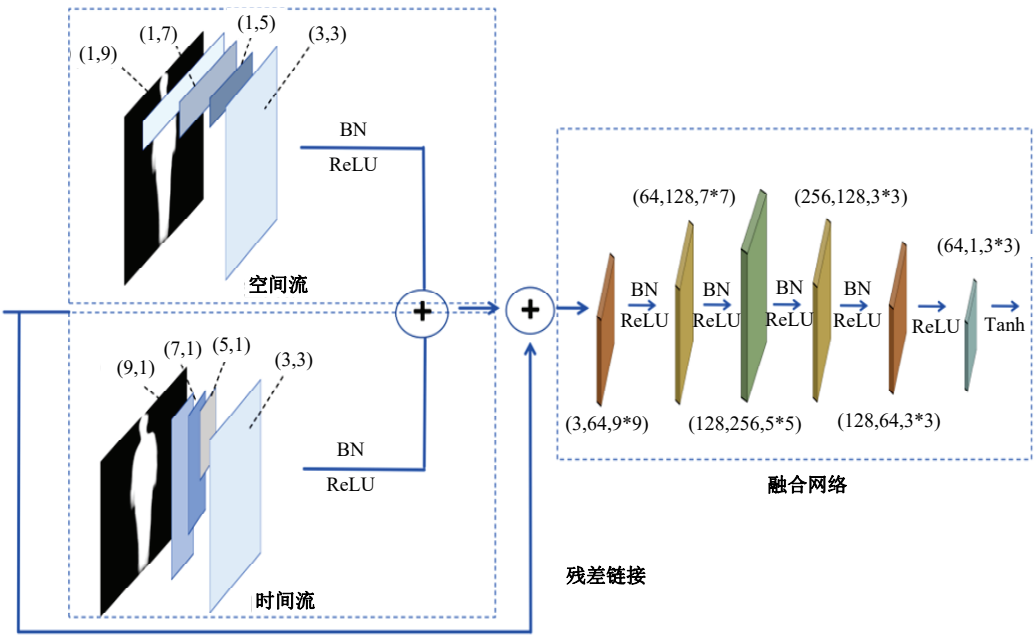


图4 DSCGAN-SE生成器编码器网络结构

Fig. 4 Network architecture of DSCGAN-SE generator

表 1 生成器网络结构参数设置

Table 1 Generator network architecture parameter settings

区域划分	层类型	卷积核	步长	深度	归一化	激活函数
空间层	卷积层	1×9	1	128	—	ReLU
空间层	卷积层	1×7	1	128	—	ReLU
空间层	卷积层	1×5	1	128	—	ReLU
空间层	卷积层	3×3	1	1	BN	ReLU
特征连接	全连接层	—	—	—	—	—
时间层	卷积层	9×1	1	128	—	ReLU
时间层	卷积层	7×1	1	128	—	ReLU
时间层	卷积层	5×1	1	128	—	ReLU
时间层	卷积层	3×3	1	1	BN	ReLU
特征连接	全连接层	—	—	—	—	—
残差层	卷积层	9×9	1	64	BN	ReLU
残差层	卷积层	7×7	1	128	BN	ReLU
残差层	卷积层	5×5	1	256	BN	ReLU
残差层	卷积层	3×3	1	128	BN	ReLU
残差层	卷积层	3×3	1	64	—	ReLU
残差层	卷积层	3×3	1	1	—	Tanh

态假图像 I_{fake} 的空间细节和结构信息,使得双通道时空卷积网络能够有效地整合和重建输入特征。通过逐步上采样的反卷积方式,不仅能逐步恢复图像的空间细节和结构信息,还能有效地融合来自空间流和时间流网络的图像特征,从而生成具有高质量和准确度的输出图像。生成器 G 的网络结构参数设置如表 1 所示。

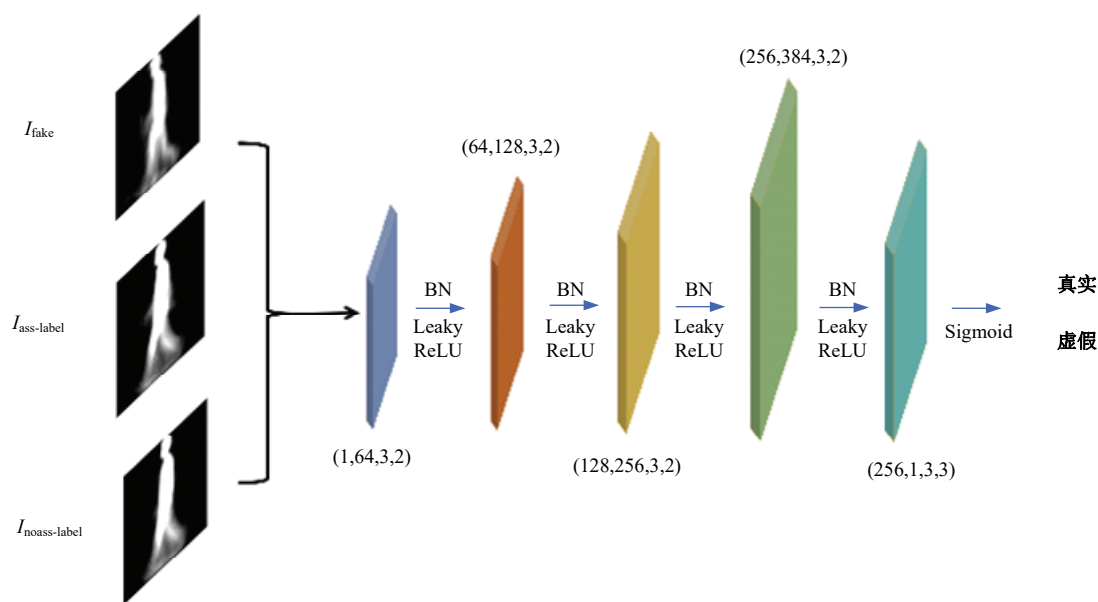
2.2 判别器网络架构

真实判别器 D_{op} 的结构图如图 5 所示,该网络首先接收生成器生成的步态假图像 I_{fake} 、相关图集 $I_{\text{ass-label}}$ 和非相关图集 $I_{\text{noass-label}}$ 作为输入,通过 5 个卷积层逐步抽取特征,每层分别配备 3×3 的卷积核,并配置 LeakyReLU 激活函数,激活函数的参数设置为 $(0.2, \text{True})$ 。真实判别器 D_{op} 可以在不同的层次上捕捉生成目标视角图像的关键特征,进而进行图像真实判别和语义信息判别。最终输出层通过 Sigmoid 函数判别图像的真伪,用以反映真实判别器 D_{op} 判断输入图像来自源图像而不是融合图像的概率。真实判别器 D_{op} 的网络结构参数设置如表 2 所示。

身份判别器 D_{tp} 的结构图如图 6 所示,与真实判别器 D_{op} 的结构类似,其通过 5 个卷积层逐步抽取特征后进行判别,每个卷积层均配备了 3×3 卷积核。然而,与真实判别器 D_{op} 不同,身份判别器 D_{tp} 除了将源图集 I_{img} 与步态假图像 I_{fake} 混合外,还将源图集 I_{img} 分别与相关图集 $I_{\text{ass-label}}$ 和非相关图集 $I_{\text{noass-label}}$ 图集混合后作为输入,强化了身份判别器 D_{tp} 对图像身份特征的判别能力。身份判别器 D_{tp} 的网络结构参数设置如表 3 所示。

2.3 风格编码器网络架构

在 DSCGAN-SE 模型中,本文首次提出风格编码器,用于从源图集 I_{img} 的步态图像和生成器生成的步态假图像 I_{fake} 中提取风格特征,其网络结

图 5 DSCGAN-SE 真实判别器 D_{op} 网络结构示意图Fig. 5 Schematic diagram of the real discriminator D_{op} network architecture in DSCGAN-SE表 2 真实判别器 D_{op} 网络结构参数设置Table 2 Real discriminator D_{op} network architecture parameter settings

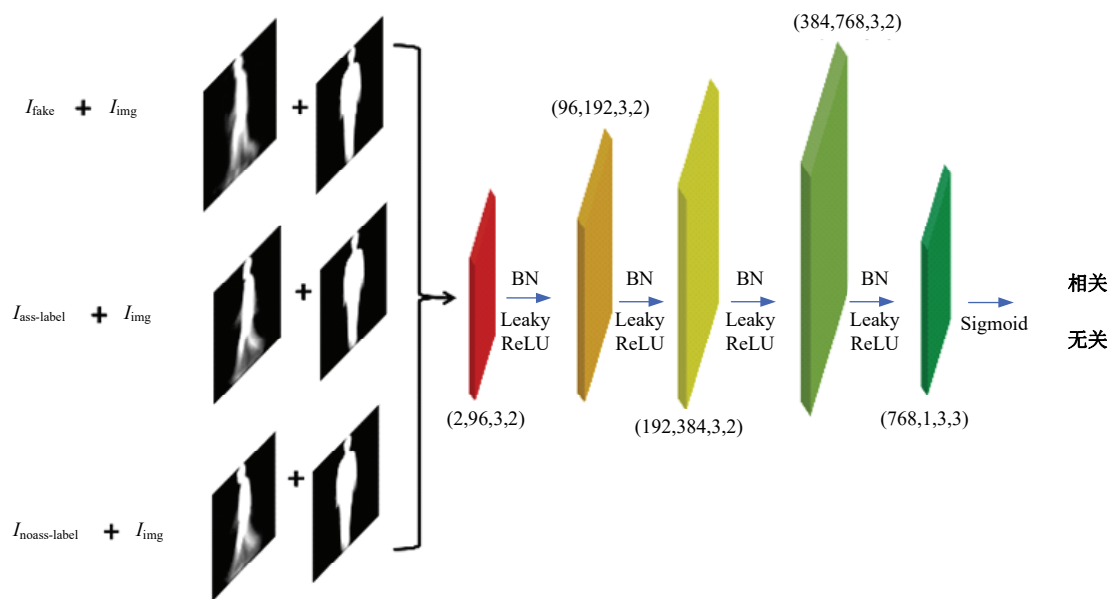
层类型	卷积核	步长	深度	归一化	激活函数
卷积层	3×3	2	64	BN	LeakyReLU
卷积层	3×3	2	128	BN	LeakyReLU
卷积层	3×3	2	256	BN	LeakyReLU
卷积层	3×3	2	384	BN	LeakyReLU
输出层	3×3	3	1	—	Sigmoid

构包括 6 个卷积层和 1 个全连接层。第 1 层卷积 (3×3 卷积核) 用于提取输入图像基本的纹理和边缘信息; 第 2 层卷积 (3×3 卷积核) 进一步捕捉

表 3 身份判别器 D_{ip} 网络结构参数设置Table 3 Identity discriminator D_{ip} network architecture parameter settings

层类型	卷积核	步长	深度	归一化	激活函数
卷积层	3×3	2	96	BN	LeakyReLU
卷积层	3×3	2	192	BN	LeakyReLU
卷积层	3×3	2	384	BN	LeakyReLU
卷积层	3×3	2	768	BN	LeakyReLU
输出层	3×3	3	1	—	Sigmoid

复杂的图像特征; 第 3 层卷积 (3×3) 加深风格编码网络特征提取能力, 以获得更丰富的图像细节。其网络结构图如图 7 所示。

图 6 DSCGAN-SE 身份判别器 D_{ip} 网络结构示意图Fig. 6 Schematic diagram of the identity discriminator D_{ip} network architecture in DSCGAN-SE

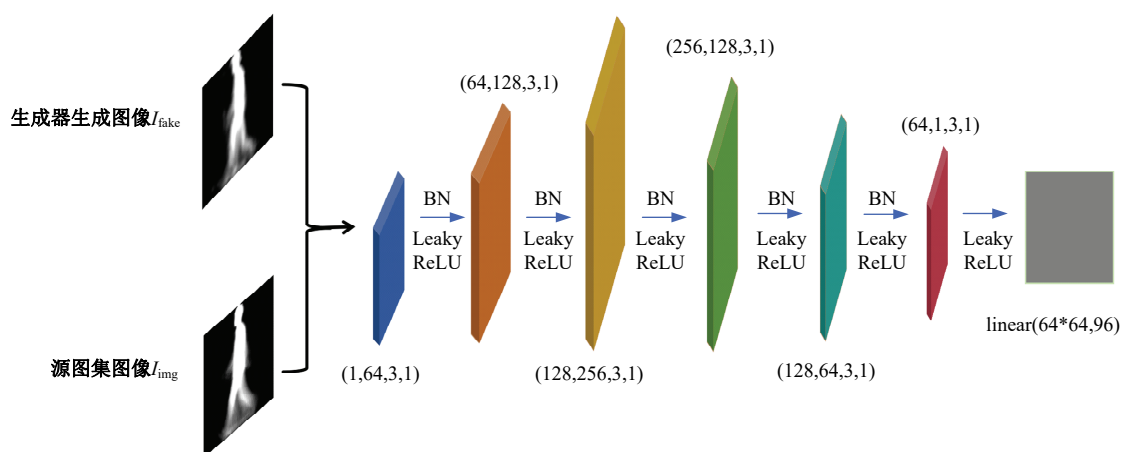


图 7 风格编码器网络结构示意图

Fig. 7 Schematic diagram of the style encoder network architecture

随后,通过3层卷积进行特征的恢复。特征图通过展平操作后,输入到全连接层以生成风格向量,该向量具有指定的风格维度。整体架构通过 LeakyReLU 激活函数和批归一化层确保网络的训练稳定性和特征的有效提取,从而精确地捕捉图像的风格信息。风格编码器是通过计算真实图像与生成图像之间的风格特征差判断生成图像与真实图像的风格差异,其中,特征差越小,表明生成图像与真实图像在风格上的一致性越高,将差异损失进行反馈达到迭代更新的目的。风格编码器的网络结构参数设置如表4所示。

表 4 风格编码器网络结构参数设置

Table 4 Style encoder network architecture parameter settings

层类型	卷积核	步长	深度	归一化	激活函数
卷积层	3×3	1	64	BN	LeakyReLU
卷积层	3×3	1	128	BN	LeakyReLU
卷积层	3×3	1	256	BN	LeakyReLU
卷积层	3×3	1	128	BN	LeakyReLU
卷积层	3×3	1	64	BN	LeakyReLU
卷积层	3×3	1	1	—	LeakyReLU
全连接层	—	—	—	—	—

2.4 损失反馈机制

DSCGAN-SE 可以通过最小化以下目标函数来优化:

$$\min_{G,D} L(G,D) = L_{adv} + \lambda_{tp} L_{tp} + \lambda_{sc} L_{sc} \quad (3)$$

式中: L_{adv} 为真实判别器对抗损失; L_{tp} 用于衡量生成步态假图像 I_{fake} 与真实图像 I_{real} 之间的差异; L_{sc} 为风格编码差异损失,用于提升生成图像的真实感; λ_{tp} 为内容感知损失权重; λ_{sc} 为风格损失权重。

生成器 G 的目标是最小化这一目标函数,而判别器 D 则尽可能将其最大化。 λ 则优化过程中不同损失函数重要性的超参数。DSCGAN-SE 利用各种损失函数提升生成图像的质量和真实性,通过调整超参数 λ 灵活控制不同损失在模型训练过程中的影响力度,从而有效平衡生成质量和训练的稳定性。在本研究中,初始参数设置为: $\lambda_{tp}=1, \lambda_{sc}=1$ 。

(1) 真实判别器 D_{op} 损失

当生成器 G 产生生成的步态假图像 I_{fake} 时,真实判别器 D_{op} 需要判断 I_{fake} 是真实还是伪造的,而生成器 G 则尝试欺骗真实判别器 D_{op} ,从而不断训练 DSCGAN-SE 模型的生成和判别能力。这一目标可以通过如下对抗损失函数 L_{adv} 实现:

$$L_{adv} = E_{(I_{ass-label}, I_{noass-label})} \times \left[\frac{1}{N} \sum_{i=1}^N \log D_{op}(I_{ass-label}, I_{noass-label}) \right] + E_{(I_{ass-label}, I_{noass-label}, I_{fake})} \times \left[\frac{1}{M} \sum_{j=1}^M \log (1 - D_{op}(G(I_{ass-label}, I_{noass-label}, I_{fake}))) \right] \quad (4)$$

式中: N 为真实图像的批量大小; $E(\cdot)$ 表示对不同数据样本的损失进行期望计算; $(I_{ass-label}, I_{noass-label})$ 为真实图像; $D_{op}(I_{ass-label}, I_{noass-label})$ 为真实判别器 D_{op} 对真实图像的预测概率; M 为生成图像的批量大小; $G(I_{ass-label}, I_{noass-label}, I_{fake})$ 为生成器 G 基于相关图集 $I_{ass-label}$ 和非相关 $I_{noass-label}$ 图集生成步态、假图像 I_{fake} 的过程; $D_{op}(G(I_{ass-label}, I_{noass-label}, I_{fake}))$ 为真实判别器 D_{op} 对生成图像的预测概率。

为了生成尽可能真实的样本,生成器 G 试图

最小化这一损失函数,而真实判别器 D_{op} 则试图最大化此损失函数。

(2) 身份判别器 D_{ip} 损失

在对抗训练中,身份判别器 D_{ip} 利用二分类交叉熵损失函数判别相关图集 $I_{ass-label}$ 、非相关 $I_{noass-label}$ 图集和生成器 G 生成的步态假图像 I_{fake} 之间的相关性。其定义如下:

$$L_{ip} = -\left[\frac{1}{N} \sum_{i=1}^N \log D_{ip}(I_{ass-label}, I_{fake}) + \frac{1}{M} \sum_{j=1}^M \log (1 - D_{ip}(G(I_{noass-label}, I_{fake}))) \right] \quad (5)$$

式中: N 为真实图像的批量大小; $(I_{ass-label}, I_{fake})$ 为相关图集与假图像的混合; $D_{ip}(I_{ass-label}, I_{fake})$ 是身份判别器 D_{ip} 对相关图集与假图像之间是否具有相关性的预测; $G(I_{noass-label}, I_{fake})$ 为生成器生成的步态假图像; $D_{ip}(G(I_{noass-label}, I_{fake}))$ 为身份判别器 D_{ip} 对非相关图集与生成图像之间相关性预测的输出。

身份判别器 D_{ip} 损失衡量了 DSCGAN-SE 模型中假图像与真图像之间的相关。它引导模型在分类任务中逐步优化,使模型能够更准确地预测输入样本的类别以及有效区分真实图像与伪图像,从而提高生成对抗网络的训练质量和生成样本的真实性。通过最小化这一损失函数,身份判别器 D_{ip} 能够在对抗训练过程中推动生成器 G 生成更具真实性的图像。

(3) 风格编码损失

DSCGAN-SE 中引入风格编码差异损失函数,旨在量化生成图像 I_{fake} 与真实图像 I_{real} 之间的风格偏差。风格编码器采用 $L1$ 损失在像素层面将生成器 G 生成的步态图像 I_{fake} 和真实图像 I_{real} 转换为高维特征向量,计算两者在风格特征空间中的绝对差异,并对所有特征的差异取平均,其定义为:

$$L_{sc} = \frac{1}{N} \sum_{i=1}^N \|S_{fake}[G(I_{real}, I_{fake})] - S_{real}[I_{real}]\|_1 \quad (6)$$

式中: $G(I_{real}, I_{fake})$ 为生成器通过输入图像 I_{real} 生成步态假图像 I_{fake} 的过程; S_{fake} 为生成图像经过风格编码器提取出的风格特征向量; S_{real} 为真实图像经过同一风格编码器提取出的风格特征向量; N 为风格特征向量的维度,即编码空间中的元素个数。

通过迭代更新 L_{sc} ,使得 DSCGAN-SE 模型在训练过程中能够优化生成图像的风格特征,使其尽可能接近真实图像的风格,从而提升生成图像的真实性和风格一致性。

3 实验与结果分析

3.1 CASIA-B 数据集

本文中采用的 CASIA-B 数据集^[13]由中国科学院自动化研究所于 2005 年收集,专为跨视角步态识别构建,包含 124 名行人,提供了每名行人的步态视频、步态轮廓图像和步态能量图像,本文选用步态能量图像(Gait energy image, GEI)作为 DSCGAN-SE 模型的输入。

实验中,数据集的划分参考 GaitSet^[7]模型中的中训练集(Medium-sample training, MT)划分方式。MT 方法将前 62 名行人的数据作为训练集,剩余 62 名行人的数据用于测试。测试时,前 4 个正常行走状态(NM#1-4)的数据用作检索库(gallery),其余数据(NM#5-6, BG#1-2, CL#1-2)根据状态划分为 3 个待检索数据(probe)进行评估。此外,本研究特别对比了包含与排除相同视角的两种情况,以更准确地展示 DSCGAN-SE 模型的跨视角识别性能。

3.2 OU-MVLP 数据集

OU-MVLP (Multi-view large population dataset)^[14]是目前全球规模最大的步态识别跨视角数据集,包含 10 307 个样本,覆盖 14 个视角($0^\circ, 15^\circ, \dots, 270^\circ$),每个视角下均包含两个步态序列(#00 和 #01),且步行状态保持一致。数据集按照官方设定分为训练集和测试集,分别包括 5 153 名样本和 5 154 名样本。在训练和测试环节中,序列 #01 和 #00 均用作训练集和测试集,为跨视角步态识别的性能评估提供了规范化的实验基准。

3.3 步态数据集图像预处理

步态能量图像(GEI)是在时间序列上平均步态轮廓图的一种表示形式,它结合了步态序列中的静态和动态信息。GEI 通过对一个步态周期内所有步态轮廓图像进行平均计算,生成步态能量图^[14]。灰度图 GEI 的定义为:

$$G(x, y) = \frac{1}{N_o} \sum_{t=1}^{N_o} S_t(x, y) \quad (7)$$

式中: N_o 为一个步态周期中目标视角步态图像的总帧数; $S_t(x, y)$ 为一个步态周期中 t 时刻的二值步态图像序列; t 为步态周期中的第 t 帧。

一个步态周期内的所有目标步态图像合成能量图的过程如图 8 所示。

3.4 评价指标及实验设置

本文使用 K 最邻近法(K -nearest neighbors,

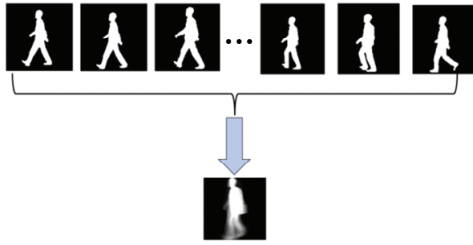


图 8 合成步态能量图 GEI

Fig. 8 Synthetic gait energy image (GEI)

KNN)作为分类器来评估模型性能。DSCGAN-SE通过将生成器 G 生成的步态图像 I_{fake} 与数据集中已知类别的真实图像 I_{real} 进行比较,KNN能够有效地衡量生成图像 I_{fake} 在特征空间中的分布与真实图像 I_{real} 的相似程度,模型通过计算 I_{fake} 与 I_{real} 中的所有样本之间的欧氏距离,找到与其距离最接近的 K 个样本。具体来说,KNN评价指标基于以下公式进行计算:

$$d(I_{\text{fake}}, I_{\text{fake},i}) = \sqrt{\sum_{j=1}^n (I_{\text{fake},j} - I_{\text{fake},i,j})^2} \quad (8)$$

式中: $I_{\text{fake},i}$ 为训练集中的样本; $I_{\text{fake},j}$ 和 $I_{\text{fake},i,j}$ 分别为样本 I_{fake} 和 $I_{\text{fake},i}$ 的第 j 个特征值; n 为特征空间维度。

找到 K 距离最近的样本后,KNN通过对这些样本的标签进行投票,选择出现次数最多的标签作为 I_{fake} 的预测标签。即:

$$\hat{y}(I_{\text{fake}}) = \text{mod } e \{ y(I_{\text{fake},i1}), y(I_{\text{fake},i2}), \dots, y(I_{\text{fake},in}) \} \quad (9)$$

式中: $y(I_{ij})$ 为第 j 个最近邻样本的标签; $\hat{y}(I_{\text{fake}})$ 为生成器 G 生成的步态图像 I_{fake} 的预测标签。

基于此,KNN能有效评估生成图像的质量,其分类准确率反映了生成图像在特征空间中与真实样本的相似程度,从而间接衡量生成器的生成质量和多样性。

3.5 实验参数设置

本实验基于PyTorch深度学习框架的环境,使用配置为AMD 5900X CPU、NVIDIA 4090显卡和32GB DDR4 3600内存的服务器。二维卷积层和反卷积层初始化权重参数是均值为0、标准差为0.02的正态分布,偏置参数设为0。使用Adam优化器进行网络参数的更新,学习率设置为0.0002/3。

数据集CASIA-B被划分为两部分。前62名人用于训练,后62名人用于测试。GEI(步态能量图)的输入和输出图像尺寸设定为 64×64 像

素,批量大小为128。DSCGAN-SE模型训练总共进行10 000次迭代。

3.6 实验结果与分析

本节将评估DSCGAN-SE在跨视角步态识别任务的实验结果,并分析了3种情况下不同视角对识别精度的影响。

为验证DSCGAN-SE模型良好的泛化能力,本节展示了基于CASIA-B数据集包含和排除相同视角3种情况下各个角度的识别精度,在图9和图10中以柱状图形式进行展示(彩图见电子版,后文同),具体3种情况下各个角度识别准确率如表5~表7所示。

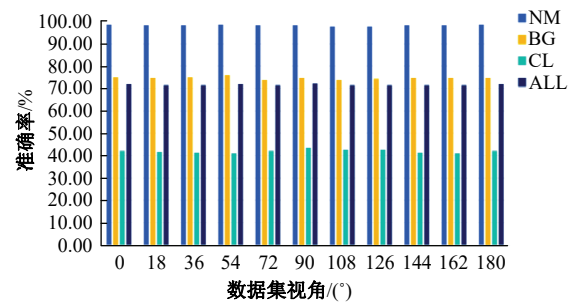


图 9 CASIA-B数据集包含相同视角3种情况下各个角度的识别精度

Fig. 9 Recognition accuracy at various angles for three cases with the same viewpoint on CASIA-B dataset

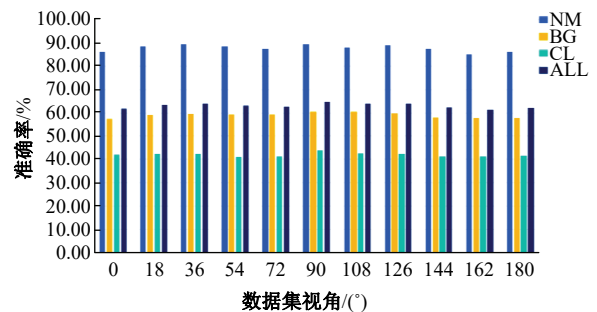


图 10 CASIA-B数据集排除相同视角3种情况下各个角度的识别精度

Fig. 10 Recognition accuracy at various angles for three cases excluding the same perspective on CASIA-B dataset

结果表明,在不同步态条件下识别率存在显著差异,其中正常行走状态下的识别率最高,而携带包和穿着外套状态下的识别率则显著降低。这表明穿着外套和携带包会影响步态特征的表达,而视角差异则会进一步降低识别准确率。视角差异造成的数据分布不一致以及对称视角(特别是在 90° 附近)的性能提升,这是由于对称视角下步

表 5 CASIA-B数据集在正常条件下的跨视角识别准确率

Table 5 Cross-perspective recognition accuracy of the CASIA-B dataset under normal conditions

数据库视角/(°)	检索库中 11 个视角	
	包含相同视角	排除相同视角
0	0.98	0.86
18	0.98	0.88
36	0.98	0.89
54	0.98	0.88
72	0.98	0.87
90	0.98	0.89
108	0.97	0.87
126	0.97	0.88
144	0.97	0.87
162	0.98	0.85
180	0.98	0.86

表 6 CASIA-B数据集在携带包下的跨视角识别准确率

Table 6 Cross-perspective recognition accuracy of the CASIA-B dataset with a bag

数据库视角/(°)	检索库中 11 个视角	
	包含相同视角	排除相同视角
0	0.75	0.57
18	0.74	0.58
36	0.75	0.59
54	0.76	0.59
72	0.73	0.59
90	0.74	0.60
108	0.73	0.60
126	0.74	0.59
144	0.74	0.57
162	0.74	0.57
180	0.74	0.57

态数据的镜像效应以及携带包和外套对人体轮廓的遮挡所致。

图 11 为 OU-MVLP 数据集在排除相同视角 3 种情况下各个角度的识别精度。从图 11 中可以看出,本文 DSCGAN-SE 模型在跨视角的大规模步态识别任务中表现出的性能优势。这种性能的提升得益于 OU-MVLP 数据集在样本数量和视角多样性上的广泛覆盖,同时该数据集避免了过多协变量的干扰,使得 DSCGAN-SE 模型成功地在各个视角下实现了更高的识别精度,充分验证了其在大规模跨视角步态识别任务中的鲁棒性和优越性能。具体各角度识别准确率如表 8 所示。

3.7 消融实验

为评估 GaitGAN 模型在引入关键模块后的

表 7 CASIA-B数据集在穿着外套下的跨视角识别准确率

Table 7 Cross-perspective recognition accuracy of the CASIA-B dataset under coat wearing conditions

数据库视角/(°)	检索库中 11 个视角	
	包含相同视角	排除相同视角
0	0.42	0.41
18	0.41	0.42
36	0.41	0.42
54	0.41	0.41
72	0.42	0.41
90	0.44	0.43
108	0.42	0.42
126	0.42	0.42
144	0.41	0.41
162	0.41	0.41
180	0.41	0.41

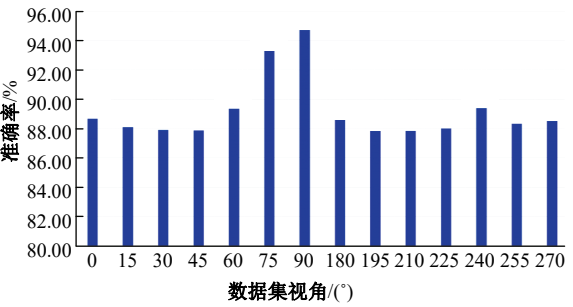


图 11 OU-MVLP数据集排除相同视角三种情况下各个角度的识别精度

Fig. 11 Recognition accuracy at various angles for three cases excluding the same perspective on OU-MVLP Dataset

性能提升,本实验通过逐步加入双通道时空卷积网络和风格编码器,验证各模块对模型性能的贡献。具体结果如表 9 所示。通过引入双通道时空卷积网络后,生成器能够同时提取步态图像的时间和空间特征,对横向时序信息和纵向空间信息的提取与融合使模型的平均识别精度从 64% 提升至 77%。进一步加入风格编码器,通过提取生成图像与真实图像的风格特征向量并计算其风格空间差异,有效优化了生成图像的真实感和语义一致性,平均识别精度提升至 81%。双通道时空卷积网络和风格编码器分别在步态特征建模和生成图像质量优化方面发挥了关键作用,进一步提升了 DSCGAN-SE 模型的跨视角步态识别能力。

3.8 实验结果可视化

目前,基于 GAN 的步态识别跨视角研究中

表 8 OU-MVLP数据集排除相同视角情况下
各个角度的识别精度

Table 8 Recognition accuracy at various angles in the
OU-MVLP dataset excluding identical
viewpoints

数据库视角/(°)	检索库中 14 个视角	
	包含相同视角	排除相同视角
0	0.88	0.87
15	0.88	0.87
30	0.87	0.86
45	0.87	0.86
60	0.89	0.88
75	0.93	0.92
90	0.94	0.94
180	0.88	0.87
195	0.87	0.86
210	0.87	0.86
225	0.88	0.86
240	0.89	0.87
255	0.88	0.87
270	0.88	0.88

表 9 双通道时空卷积和风格编码器对跨视角识别
准确率的影响(基于 CASIA-B 数据集)

Table 9 Impact of dual-channel spatiotemporal
convolution and style encoder on
cross-view recognition accuracy
based on the CASIA-B dataset

网络模型	验证角度			平均值
	54°	90°	126°	
GaitGAN	0.65	0.61	0.66	0.64
GaitGAN+双通道时空 卷积网络+风格编码器	0.88	0.89	0.88	0.88

都需要实现将原本视角与目标进行评估之后,才能生成特定的步态图像,如果需要进行多视角的步态图像转换,则需要构建多个模型。

本文提出的 DSCGAN-SE 模型可将步态模板从任意视角转换成目标视角,图 12 为 DSCGAN-SE 利用 CASIA-B 数据集随机抽出 128 个行人进行 90°典型视角的跨视角转换可视化,每三行为一组,其中第 1 行为生成器生成的目标视角图像,第 2 行为数据集中真实目标视角图像,第 3 行为原始步态图像。可以看出,本文训练 DSCGAN-SE 模型即使在视角变化较大情况下,生成器生成的步态图像也与真实的目标视角的步态图像难以区分。图 13 为 DSCGAN-SE 利用 OU-MVLP 数据集随机抽出 128 个行人进行 90°

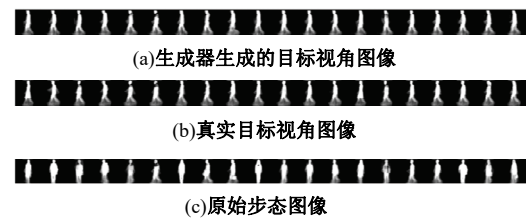


图 12 DSCGAN-SE 模型原本视角步态图像与
目标视角步态图像转化的可视化结果
(基于 CASIA-B 数据集)

Fig. 12 Visualization results of DSCGAN-SE model's
original viewpoint and target viewpoint gait
image transformations based on the
CASIA-B dataset

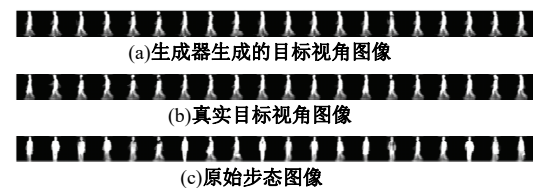


图 13 DSCGAN-SE 模型原本视角步态图像与
目标视角步态图像转化的可视化结果
(基于 OU-MVLP 数据集)

Fig. 13 Visualization results of DSCGAN-SE model's
original viewpoint and target viewpoint gait
image transformations based on the
OU-MVLP dataset

典型视角的跨视角转换可视化。

3.9 CASIA-B 数据集上跨视角步态识别对比

为了证明 DSCGAN-SE 模型的泛化性,本文将 DSCGAN-SE 与其他基于 GAN 的步态识别方法(GaitGAN^[12]、TCC-GAN^[15]、SE-CLGait^[16]、Dtrain-GAN^[17]、SA-GAN^[18]) 在 CASIA-B 数据集正常行走条件下的步态识别精度进行对比分析,结果如表 10 所示。

从表 10 中可以看出,与其他方法相比,本文 DSCGAN-SE 模型的性能表现最佳。GaitGAN^[12]虽在识别率上取得了一定提升,但由于过度依赖于 90°侧视图为标准进行特征匹配,忽视了中间层和时序特征的提取,导致在多视角条件下生成图像的质量较为有限。SE-GLGait^[16]中使用的 SPTEM 和 LSGF 方法都是基于步态视频序列生成,模型在处理其他类型的输入(如静态图像)时可能不适用。SA-GAN^[18]使用的自注意力机制虽然在全局特征提取上效果出色,但可能会忽略一些步态图像中的细微局部特征,导致模型难以捕捉不同步态的细微差异。Dtrain-GAN^[17]利

表 10 正常行走下的平均识别率对比
(基于 CASIA-B 数据集)

Table 10 Comparison of average recognition rates during normal walking (based on the CASIA-B dataset)

模型	验证角度			
	54°	90°	126°	平均
GaitGAN	0.65	0.61	0.66	0.64
TCC-GAN	0.81	0.73	0.79	0.78
SE-CLGait	0.76	0.76	0.69	0.73
Dtrain-GAN	0.64	0.61	0.68	0.64
SA-GAN	0.81	0.72	0.80	0.77
本文	0.88	0.89	0.88	0.88

用三元组损失增大不同类步态特征间的距离,并缩小同类特征间的距离,但其依赖于精确的负样本选择,且计算复杂度高,对训练时间和资源需求较大。相比之下,DSCGAN-SE不仅在全局和局部特征提取上进行了优化,还通过改进 TCC-GAN^[15]的双通道卷积网络生成器设计,进一步增强了模型的时空特征提取能力。

DSCGAN-SE 模型能够实现从任意视角到目标视角的步态模板转换,关键在于双通道时空卷积网络结构。该结构通过并行处理空间和时序特征,增强了模型对步态图像时序动态和空间结构的理解,提升了生成器在多视角条件下的表现能力。这种特征分离的设计让模型能够更好地捕捉跨视角生成任务中的复杂细节与变化。此外,为进一步提高生成图像的质量,DSCGAN-SE 还引入了风格编码器,确保生成图像在语义层面的一致性,从而提升生成器的稳定性和鲁棒性。综合这些创新性设计,DSCGAN-SE 在处理跨视角步态识别任务时展现了卓越的性能,成功实现了步态模板的高质量生成和视角转换。

相比于基准模型 GaitGAN^[14],DSCGAN-SE 模型在 54°、90° 和 126° 三个视角下的识别率分别提高了 23%、28% 和 22%。这说明 DSCGAN-SE 模型不仅很好地解决了 GaitGAN 模型中基础 GAN 结构性能有限问题,同时还改进了 TCC-GAN 中将双通道卷积网络运用在生成器设计中的方法,因而在跨视角步态识别中表现出色。

3.10 OU-MVLP 数据集上跨视角步态识别对比

本文提出的 DSCGAN-SE 模型不仅在

CASIA-B 数据集上表现出色,在 OU-MVLP 数据集上的实验结果同样优异。如表 11 所示,DSCGAN-SE 在 14 个视角的识别准确率较优于 GEINet^[19]、GaitSet^[7]、GaitGANv2^[20]、MvG-GAN^[21]、FSEAM^[22]和 TSGM^[23]模型。该模型能够在复杂视角变化条件下精准提取步态特征,并在大规模数据集上的应用中展现出卓越的泛化能力。其性能优势得益于在生成对抗网络框架中融合了风格编码和双通道时空卷积网络,有效增强了特征表达能力和跨视角匹配性能,从而实现更高的识别精度。实验结果充分证明了 DSCGAN-SE 在跨视角步态识别任务中的有效性和鲁棒性,为步态识别技术的实际应用提供了重要支持。

表 11 各方法在 OU-MVLP 数据集上的识别准确率对比

Table 11 Comparison of recognition accuracy of various methods on the OU-MVLP dataset

视角/ (°)	基于全部 14 个视角的注册库						本文 模型
	GEI- Net	MvG- GAN	Gait- Set	Gait- GAN2	FSEAM	TSGM	
0	0.11	0.52	0.79	0.81	0.81	0.64	0.87
15	0.29	0.62	0.87	0.88	0.90	0.79	0.87
30	0.41	0.63	0.89	0.90	0.89	0.87	0.86
45	0.45	0.57	0.90	0.90	0.90	0.88	0.86
60	0.39	0.55	0.88	0.88	0.89	0.80	0.88
75	0.41	0.64	0.88	0.89	0.89	0.84	0.92
90	0.38	0.61	0.87	0.88	0.88	0.81	0.94
180	0.14	0.54	0.81	0.83	0.83	0.65	0.87
195	0.33	0.58	0.86	0.87	0.88	0.78	0.86
210	0.43	0.59	0.89	0.89	0.89	0.85	0.86
225	0.45	0.58	0.89	0.89	0.89	0.87	0.86
240	0.39	0.56	0.87	0.87	0.88	0.79	0.86
255	0.40	0.57	0.87	0.88	0.88	0.83	0.87
270	0.36	0.56	0.86	0.86	0.86	0.79	0.87
平均值	0.35	0.58	0.87	0.87	0.88	0.80	0.88

4 结束语

为克服 GaitGAN 模型对视角依赖性高和动态变化适应性差的问题,同时为提高跨视角步态识别的准确率,本文提出了一种基于 DSCGAN-SE 模型的步态识别方法。该模型在 GaitGAN 的基础上引入了双通道时空卷积和风格编码,解决了 GaitGAN 模型对视角依赖性高和动态变化适应性差的问题,进一步提高了跨视角步态识别的准确率,与基准 GaitGAN 模型相比,本文模型在

54°、90°和126°三个具有代表性的步态视角下的识别率分别提升了23%、28%和22%。

参考文献:

- [1] 李贻斌, 郭佳旻, 张勤. 人体步态识别方法与技术[J]. 吉林大学学报: 工学版, 2020, 50(1): 1-18.
Li Yi-bin, Guo Jia-min, Zhang Qin. Methods and technologies for human gait recognition[J]. Journal of Jilin University (Engineering and Technology Edition), 2020, 50(1): 1-18.
- [2] Sepas-Moghaddam A, Etemad A. Deep gait recognition: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 264-284.
- [3] 段成阁, 刘康康, 李福全. 步态识别技术综述[J]. 中国人民公安大学学报: 自然科学版, 2022, 28(4): 75-80.
Duan Cheng-ge, Liu Kang-kang, Li Fu-quan. Overview of gait recognition technology[J]. Journal of People's Public Security University of China (Natural Science Edition), 2022, 28(4): 75-80.
- [4] 刘晓芳, 周航, 韩权, 等. 基于视觉的步态识别研究综述[J]. 小型微型计算机系统, 2018, 39(8): 1685-1692.
Liu Xiao-fang, Zhou Hang, Han Quan, et al. Review of vision-based gait recognition research[J]. Mini-Micro Systems, 2018, 39(8): 1685-1692.
- [5] 刘相锋, 周航. 步态识别技术及其在视频监控中的应用[J]. 电视技术, 2011, 35(1): 119-121.
Liu Xiang-feng, Zhou Hang. Gait recognition technology and its application in video surveillance[J]. Television Technology, 2011, 35(1): 119-121.
- [6] 陈福仕, 沈尧, 周池春, 等. 无监督学习步态识别综述[J]. 计算机科学与探索, 2024, 18(8): 2014-2033.
Chen Fu-shi, Shen Yao, Zhou Chi-chun, et al. A review of unsupervised learning for gait recognition[J]. Computer Science and Exploration, 2024, 18(8): 2014-2033.
- [7] 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, 44(7): 3467-3478.
- [8] Limcharoen P, Khamsemanan N, Nattee C. Gait recognition and re-identification based on regional LSTM for 2-second walks[J]. IEEE Access, 2021, 9: 112057-112068.
- [9] Goodfellow I, Pougetabadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in Neural Information Processing Systems, Quebec, Canada, 2014: 2672-2680.
- [10] Aljazeera M A A, Makihara Y, Muramatsu D, et al. Batch hard contrastive loss and its application to cross-view gait recognition[J]. IEEE Access, 2023, 11: 31177-31187.
- [11] Wang L H, Xiao P F, Zhang X L, et al. A fine-grained unsupervised domain adaptation framework for semantic segmentation of remote sensing images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2026, 16: 4109-4121.
- [12] Yu S, Chen H E, Reyes B G, et al. GaitGAN: invariant gait feature extraction using generative adversarial networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 2017: 532-539.
- [13] Yu S Q, Tan D L, Tan T N. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition[C]//Proc of the 18th International Conference on Pattern Recognition, Washington, USA, 2006: 441-444.
- [14] Xu C, Makihara Y, Liao R, et al. Applications: real-time gait-based age estimation and gender classification from a single image[C]//Proceedings of the 2021 IEEE/CVF Winter Conference on Applications of Computer Vision, online, 2021: 3459-3469.
- [15] 王宇, 夏懿. 基于双通道循环一致性GAN的跨视角步态识别研究[J]. 计算机应用研究, 2022, 39(1): 259-264.
Wang Yu, Xia Yi. Cross-view gait recognition based on dual-channel cycle-consistent GAN[J]. Computer Applications and Research, 2022, 39(1): 259-264.
- [16] 张云佐, 董旭. 基于双分支时空步态特征融合的深度学习步态识别[J]. 控制与决策, 2024, 39(4): 1403-1408.
Zhang Yun-zhuo, Dong Xu. Gait recognition based on dual-branch spatiotemporal feature fusion using deep learning[J]. Control and Decision, 2024, 39(4): 1403-1408.
- [17] 李凯, 曹可凡, 沈皓凝. 基于步态序列的跨视角步态识别[J]. 河北大学学报: 自然科学版, 2021, 41(3): 311-320.
Li Kai, Cao Ke-fan, Shen Hao-ning. Cross-view gait recognition based on gait sequences[J]. Journal of

- Hebei University (Natural Science Edition), 2021, 41(3): 311-320.
- [18] 张红颖, 包雯静. 融合自注意力机制的生成对抗网络跨视角步态识别[J]. 中国图象图形学报, 2022, 27(4): 1097-1109.
- Zhang Hong-ying, Bao Wen-jing. Cross-view gait recognition using generative adversarial networks with self-attention mechanism[J]. Journal of Image and Graphics, 2022, 27(4): 1097-1109.
- [19] Shiraga K, Makihara Y, Muramatsu D, et al. GEI-Net: view-invariant gait recognition using a convolutional neural network[C]//Proceedings of the 2016 International Conference on Biometrics. Halmstad: Sweden, 2016: 1-8.
- [20] Yu S, Liao R, An W, et al. GaitGANv2: Invariant gait feature extraction using generative adversarial networks[J]. Pattern Recognition, 2019, 87: 179-189.
- [21] Chen X, Luo X, Weng J, et al. Multi-view gait image generation for cross-view gait recognition[J]. IEEE Transactions on Image Processing, 2021, 30: 3041-3055.
- [22] 陈万志, 唐浩博, 王天元. 融合轮廓增强和注意力机制的改进GaitSet步态识别方法[J]. 电子测量与仪器学报, 2024, 38(1): 203-210.
- Chen Wan-zhi, Tang Hao-bo, Wang Tian-yuan. Improved GaitSet gait recognition method integrating contour enhancement and attention mechanism[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(1): 203-210.
- [23] 汪堃, 雷一鸣, 张军平. 基于双流步态网络的跨视角步态识别[J]. 模式识别与人工智能, 2020, 33(5): 383-392.
- Wang Kun, Lei Yi-ming, Zhang Jun-ping. Cross-view gait recognition based on dual-stream gait network[J]. Pattern Recognition and Artificial Intelligence, 2020, 33(5): 383-392.