

DOI: 10.13382/j.jemi.B2104954

# 结合残差网络与多级分块结构的步态识别方法\*

张红颖 田鹏华

(中国民航大学电子信息与自动化学院 天津 300300)

**摘要:**针对步态识别中由于衣着与背包的遮挡造成不能提取有鉴别性的步态特征,从而导致识别准确率不高的问题,提出一种结合残差网络和多级分块结构的步态识别方法。首先在水平方向上对步态能量图进行不同尺度的多级分块,以提取不同区域的细粒度特征,减少局部遮挡对于其他区域的影响,同时为了更好地学习在步态中运动频率最高区域的特征,在腿部加入Inception模块;其次为了提升网络模型的识别精度,结合交叉熵损失、三元组损失、L2正则化对残差网络的权值进行约束。最后在公开的步态数据集CASIA-B和OU-ISIR Treadmill B上进行实验,在携带背包或不同衣着条件下的识别率分别达到了87.5%、82.6%,表明该模型对于衣着与携带背包的条件具有鲁棒性。

**关键词:**深度学习;步态识别;残差网络

**中图分类号:** TP18 **文献标识码:** B **国家标准学科分类代码:** D510.4

## Gait recognition method combining residual network and multi-level block structure

Zhang Hongying Tian Penghua

(College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

**Abstract:** In gait recognition, the discriminative gait feature cannot be extracted due to the occlusion of clothing and backpack, which leads to the low recognition accuracy. A gait recognition method combining ResNet and multi-level block structure is proposed in this paper. First of all, the gait energy map is divided into different scales in the horizontal direction to extract the fine-grained features of different regions, which reduce the impact of local occlusion on other regions. At the same time, in order to better learn the characteristics of the region with the highest motion frequency, the Inception module is added. Secondly, in order to improve the recognition accuracy of the network model, cross-entropy loss, triple loss and L2 regularization are utilized to constrain the weight of the residual network. Finally, experiments were processed in the public gait data set CASIA-B and OU-ISIR Treadmill B, and the recognition rate reached 87.5% and 82.6% under different clothing or backpack conditions. It is indicated that under these conditions, the method could obtain favorable veracity and good robustness.

**Keywords:** deep learning; gait recognition; ResNet

## 0 引言

人体步态<sup>[1]</sup>作为生物特征的一种,相比于其他生物特征,如人脸、指纹等静态特征,更难以被模仿,并且步态数据的获取具有距离远、非侵犯性等优势,对于低分辨率的步态图像也能完成识别,这些优势使得基于步态的个人身份识别在许多应用中更具吸引力,如监控系统、取证

和刑事调查<sup>[2]</sup>。然而,考虑到实际应用中的复杂场景,存在许多影响步态识别的因素,如服装变化、视角、步行速度等。特别地,人们在日常生活中可能随身携带各种物品以及穿着不同的服装,它们具有各种形状和大小,并且位于身体上的不同部位,如手上、腰间、后背等,这些变化造成的遮挡导致了鲁棒步态识别的困难,因此针对衣着与背包的步态识别成为近几年的研究热点。

当前针对衣着与背包的这类协变量的研究主要分为

收稿日期: 2021-11-29 Received Date: 2021-11-29

\* 基金项目: 国家重点研发计划(2018YFB1601200)、中国民航大学天津市智能信号与图像处理重点实验室开放基金(2019ASP-TJ06)项目资助

两大类:基于模型的方法<sup>[3]</sup>和基于轮廓的方法<sup>[4]</sup>。基于模型的方法需要高分辨率图像来获取详细的人体结构信息,而基于轮廓的方法只需要低分辨率图像就能提取有效的步态特征,具有广泛的适用性。在基于模型的方法中,步态的表示一般为一个步态周期的序列集合<sup>[5]</sup>或者类能量图(如步态能量图<sup>[6]</sup>(gait energy image, GEI)、步态流图<sup>[7]</sup>(gait flow image, GFI)、计时步态图<sup>[8]</sup>(chrono-gait image, CGI)、累积预测图像<sup>[9]</sup>(accumulated prediction image, API))。序列集合虽然包含完整的特征,但其需要的计算量太大。相比之下,类能量图不仅能有效的表示步态特征,而且其内存占用仅为序列的几十分之一,因而被广泛使用。其中的 GEI 更是被验证过最简单且高效的步态表示方法。

步态特征的学习包括传统方法和深度学习方法。深度学习由于使用卷积神经网络<sup>[10]</sup>对特征进行自主提取,在步态识别领域取得了更好的效果。其中残差网络(ResNet)作为卷积神经网络的一种改进,提高了其信息流通能力,避免了网络过深所引起的退化问题。因此对于衣着与携带背包的遮挡问题,许多研究学者基于残差网络进行了研究,如王杰等<sup>[11]</sup>提出将残差网络的思想结合到 LeNet 网络模型,针对降低协变量对步态识别的影响。Wang 等<sup>[12]</sup>提出了一种基于残差网络的特征学习和识别算法,首先利用深度卷积神经网络获取人体行走特征的矢量图像,然后通过拉格朗日方法得到动态特征向量,以表示步态拟合中特征向量之间的映射关系,提高了对衣着、携带物等多种协变量因素的鲁棒性能。刘太阳等<sup>[13]</sup>提出一种双流残差网络结构,通过引入 Set

Pooling 操作提取步态周期的时序特征,改进三元组损失函数等,提升了行人轮廓存在遮挡情况下的识别准确率。胡少晖等<sup>[14]</sup>提出多支路残差网络,利用残差学习和多尺度特征融合技术,增强网络的特征学习能力,提升了步态识别的准确率。但以上这些方法均是利用残差网络提取全局特征或者将提取到的全局特征分为不同尺度的局部特征,再将其放入分类器中进行匹配。如果以全局特征进行匹配会忽略局部特征的影响,而将全局特征划分为不同尺度的局部特征的方法虽然能突出局部信息,但受遮挡的影响较大,因为提取全局特征时并没有对遮挡进行任何处理,影响了分类效果的进一步提高。

因此,本文对残差网络进行改进,在特征提取阶段对步态特征进行水平多级分块,以逐级融合的方式来减弱因遮挡造成的局部特征的互相影响,提取有区分性的全局特征。同时为了提高腿部区域步态特征的完整性,加入 Inception 模块,最后对残差模块 BasicBlock 进行改进以提升模型训练速度。

### 1 网络模型

本文提出的步态识别网络模型如图 1 所示。输入为 GEI,网络结构分为 3 个部分:1)依据人体躯干的比例提取不同区域的原始特征,并对区域特征进行融合;2)提取步态全局特征;3)对特征进行分类,通过结合交叉熵损失、三元组损失以及 L2 正则化对网络权值进行优化以提升分类准确率。

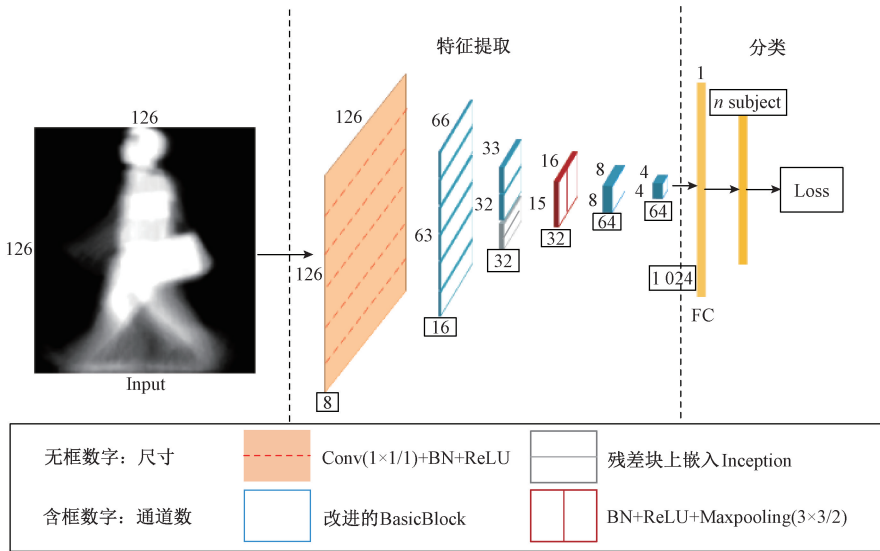


图 1 本文方法的网络框架

Fig. 1 Framework of the proposed method

在网络模型的第 1 部分中,先使用 1×1 的卷积核对原始特征空间进行升维以提供多维度的原始特征,增强

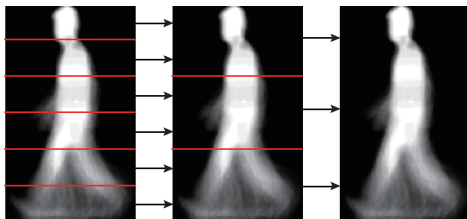
提取特征的多样性。然后在水平方向上对其进行水平多级分块:第 1 级为 6 块残差模块(头、肩、腰、大腿、小腿、

脚)、第 2 级为 3 块残差模块、第 3 级为最大池化层,以分别提取身体不同区域的局部特征,屏蔽不同区域之间特征的影响,通过逐步扩大特征提取的区域以对小区域进行逐步融合,并在第 2 级的腿部区域(脚+小腿)加入 Inception 模块进行多尺度的特征提取,最后对各区域特征进行最大值池化以融合不同区域的鉴别性特征。第 2 部分通过使用两个残差模块逐步缩小特征尺寸以提取全局特征。第 3 部分对提取到的特征进行一维展开,以交叉熵损失进行分类,利用三元组损失对难以识别的样本进行处理,扩大其类间距离,最后使用 L2 正则化对模型的权值进行平滑处理,增加模型的鲁棒性。

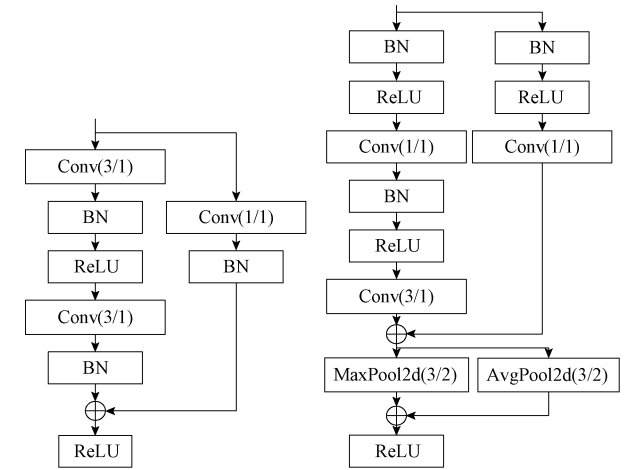
### 1.1 改进的 BasicBlock

传统残差网络的残差模块 BasicBlock 的结构如图 2 (a) 所示,其中对数据进行批量正则化(batch normalization, BN)的目的是找到一个线性和非线性之间的平衡点,以此增强非线性带来的较强表达能力,避免输出数据落在非线性区域的两端,从而减慢网络的收敛速度。但是将 BN 放在卷积层之后会使得最后求和的结果偏离平衡点,导致进入下一个残差块时,输入数据也产生偏离。因此本文对残差模块进行改进,改进后的结构如图 2(b) 所示。

首先将 BasicBlock 中卷积层之后的 BN 和 ReLU 置于卷积层之前,使得每次卷积和 ReLU 都在平衡点附近;同时,在每个残差块中只设置一个 3×3 卷积层,以此提高信息流动的密度;且在卷积之前采用 1×1 的卷积核进行升维操作,相比直接通过 3×3 卷积层进行升维能大幅度地减少参数量和计算量,优化网络的训练速度。最后在



(a) 对特征图进行多级分块  
(a) Multi-level block of feature maps



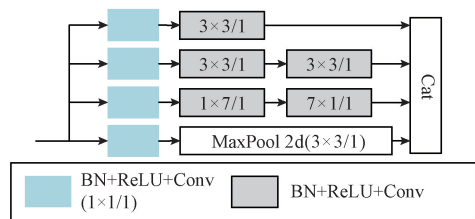
(a) 原始的 BasicBlock (a) Original BasicBlock  
(b) 改进后的 BasicBlock (b) Improved BasicBlock

图 2 对 BasicBlock 的改进  
Fig. 2 Improvements to BasicBlock

卷积层之后加入最大值池化和均值池化以选择有效特征,最大值池化能够选择代表图像纹理信息的特征,均值池化则能够选择代表图像背景信息的特征。

### 1.2 多级分块结构

衣着与背包造成的遮挡是对于步态特征局部区域的遮挡,在不同的区域有着不同的变化规则,为了有效提取局部特征,削弱局部区域之间的影响,本文根据人体在垂直方向上的比例的相似性,将步态特征进行水平多级分块,如图 3(a) 所示。



(b) 嵌入的 Inception 模块  
(b) Embedded Inception module

图 3 水平多级分块结构  
Fig. 3 Horizontal multi-level block structure

第 1 级将特征分为 6 块,为了能够有效地隔开人体的各个部分,使得能够独立地提取人体的头部、肩部、腰部、大腿部、小腿部、脚部的特征。在第 2 级将 6 块特征相邻的部分合并成 3 块,使得在步态中变化频率最快的腿部区域能够与其他区域隔开,并在该区域的残差模块中嵌入 Inception 模块以提取多尺度的更具有鉴别性的步态特征,如图 3(b) 所示,Cat 表示在通道维度上进行合

并。第 3 级使用最大值池化对第 2 级提取的局部特征进行整合,选择最具鉴别性的特征作为全局的初始特征。该方法将各部分特征之间的干扰程度大幅度降低,对于衣着和携带背包造成的干扰有一定的屏蔽作用。

### 1.3 损失函数

为了提升网络模型识别精度,本文结合交叉熵损失、L2 正则化、三元组损失对网络进行训练。交叉熵表示为

真实概率分布与预测概率分布之间的差异,通过训练能得到各个类别的基本分布,交叉熵的定义如式(1)所示。L2正则化其主要目的是控制模型复杂度,平滑不同类之间的分界线,增大网络的鲁棒性,L2正则化的定义如式(2)所示。三元组损失的优势在于细节区分,即当两个输入相似时,三元组损失能够更好地对细节进行建模,相当于加入了两个输入差异性差异的度量,学习到输入的更好表示,从而极大地提升深度特征的判别能力,三元组损失的定义如式(3)所示。

$$L_{ce}(p, q) = -\frac{1}{N} \sum_{i=1}^N p(x_i) \log(q(x_i)) \quad (1)$$

式中: $p$ 表示真实概率, $q$ 表示预测概率, $x$ 表示每一批输入图像的概率分布, $N$ 表示输入批次中样本的数量, $i$ 表示该批次的第几张图像。

$$L_{l2}(\omega) = \sum_i \omega_i^2 \quad (2)$$

式中: $\omega$ 表示网络模型中的权重值, $i$ 表示权重的序号。

$$L_{tr}(a, p, n) = \frac{1}{N} \sum_{i=1}^N \max\{d(a, p) - d(a, n) + margin, 0\} \quad (3)$$

其中, $d$ 表示欧氏距离, $a$ 表示基准样本的输出向量, $p$ 表示与 $a$ 同类但不一样的正样本的输出向量, $n$ 表示与 $a$ 不同类的负样本, $margin$ 为超参数,决定了训练后正负样本的差异程度, $N$ 表示样本数量, $i$ 表示样本序号。由于一个数据集的三元组 $\langle a, p, n \rangle$ 的组合非常多,全部进行训练会降低效率,所以通常有三类选择策略,为简单三元组、一般三元组、困难三元组,本文选取每一批输入图像的一般三元组( $d(a, p) < d(a, n) < d(a, p) + margin$ )和困难三元组( $d(a, n) < d(a, p)$ )。

综上,模型的总损失函数为:

$$Loss = \alpha L_{ce}(p, q) + \beta L_{l2}(\omega) + \gamma L_{tr}(a, p, n) \quad (4)$$

式中: $\alpha + \beta + \gamma = 1$ , $\alpha, \beta, \gamma$ 均为超参数,表示各项所占的比例。

## 2 实验与分析

本文在两个公开数据集 CASIA-B 和 OU-ISIR Treadmill B 上对提出的方法进行验证,实验在 Ubuntu18.04 系统下基于 Pytorch 1.6.0 实现,编程语言为 Python 3.7.7,计算机配置为 Intel Xeon(R) Sliver 4112 CPU 2.6 GHz 和两块 GeForce RTX 2080Ti 显卡。实验验证分为 3 个部分,第 1 部分是对改进之前和改进之后的消融性实验,第 2 部分是与近几年其他的经典基准模型进行对比,第 3 部分是对实验结果的分析。

### 2.1 数据集和训练设置

CASIA-B 数据集(YU、Tan 和 Tan 2006)是一个常用

的公开步态数据集。它包含 124 个对象、3 种步态类型和 11 个视图( $0^\circ, 18^\circ, \dots, 180^\circ$ )。步态类型包括正常行走(NM)(每个对象 6 个序列)、携带背包行走(BG)(每个对象 2 个序列)和穿着外套或夹克(CL)(每个对象 2 个序列),即每个对象有  $11 \times 10 = 110$  序列。由于该数据集的训练集和测试集(包含注册集(gallery)与验证集(probe))没有正式划分,参考了文献[15]中的设置,如表 1 所示。

表 1 CASIA-B 的实验设置

数据集	ID	视频序列
训练集	000-073	nm01-nm06, bg01-bg02, cl01-cl02
测试集-gallery	074-123	nm01-nm04 nm05-nm06
测试集-probe	074-123	bg01-bg02 cl01-cl02

OU-ISIR Treadmill B (Makihara、Mannami 和 Tsuji (2012))数据集是一个大型步态数据集,用于评估存在服装变化时的步态方法。它总共有 1 350 个视频序列(包含 68 个对象和 32 种类型的服装组合),每个对象最少为 9 种服装组合,最多有 28 种服装组合。表 2 显示了构建数据集时使用的 15 种不同类型的服装。表 3 显示了本文所用的基于 15 种不同类型服装的服装组合。整个数据集分为两部分,50 个对象(包含 881 个视频序列)用于训练模型,服装组合类型最多的 18 个对象(包含 469 个视频序列)用于测试。两个子集中的对象不重叠,以测试中所有对象的服装组合 type 9(RP+FS)作为注册集,验证集由其余的服装组合组成。

表 2 OU-ISIR treadmill B 数据集中使用的衣服列表

RP-Regular pants	BP-Baggy pants	CW-Casual wear
Sk-Skirt	CP-Casual pants	HS-Half shirt
LC-Long coat	Pk-Parka	DJ-Down jacket
SP-Short pants	Ht-Hat	FS-Full shirt
Cs-Casquette cap	RC-Rain coat	Mf-Muffler

表 3 OU-ISIR treadmill B 数据集中不同的服装组合

type	s1	s2	s3	type	s1	s2	type	s1	s2
3	RP	HS	Ht	0	CP	CW	F	CP	FS
4	RP	HS	Cs	2	RP	HS	G	CP	Pk
6	RP	LC	Mf	5	RP	LC	H	CP	DJ
7	RP	LC	Ht	9	RP	FS	I	BP	HS
8	RP	LC	Cs	A	RP	Pk	J	BP	LC
C	RP	DJ	Mf	B	RP	DJ	K	BP	FS
N	SP	HS	-	D	CP	HS	L	BP	Pk
S	Sk	HS	-	E	CP	LC	M	BP	DJ
V	Sk	DJ	-	P	SP	Pk	R	RC	-
U	Sk	PK	-	T	Sk	FS	-	-	-



在模型训练部分,本文设置数据的训练批量是 500,学习率前 5 000 轮为  $1 \times 10^{-4}$ ,后 5 000 轮为  $1 \times 10^{-5}$ ,所有卷积层的偏置(bias)为 0, padding 为 1, 优化器为 Adam, 其参数设置为(0.5, 0.999), 在损失函数上, L2 正则化的权重(weight)设置为 0.001, 三元组损失的 margin 值设置为 0.2。Loss 的超参数( $\alpha, \beta, \gamma$ )在训练的前 1 000 轮为(1, 0, 0), 中间的 4 000 轮为(0.4, 0.2, 0.4), 后 5 000 轮均为(0.6, 0.2, 0.2)。

### 2.2 在 CASIA-B 上的实验与分析

为了验证本文提出的方法的有效性,在 CASIA-B 上以(bg-01, bg-02)作为验证数据进行了消融性实验,结果如表 4 所示。

可以看出,三元组对于模型训练的提升效果最好,其次是水平多级分块特征,然后为 L2 正则化,最终对这 3

表 4 在 CASIA-B 上的消融实验

Table 4 Ablation experiment on CASIA-B

多级分块结构	三元组	Inception	L2 正则化	bg-mean/%
√				0.70
	√			0.82
√		√		0.72
			√	0.73
√	√	√		0.83
√		√	√	0.74
√	√	√	√	<b>0.87</b>

个进行结合得到了最好的准确率。

将本文提出的方法与 Wu 等<sup>[15]</sup>提出的 Deep CNN 的方法, Chao 等<sup>[16]</sup>提出的 GaitSet 模型, Zhang 等<sup>[17]</sup>提出的联合 CNN 框架进行了对比,如表 5 所示。

表 5 CASIA-B 上的平均 rank-1 准确率,不包括相同视角的情况

Table 5 Averaged rank-1 accuracy on CASIA-B, excluding identical-view cases

(%)

Gallery NM #1-4	0° ~ 180° (exclude identical viewing angle)											Mean
Probe NM #5-6	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
LB <sup>[15]</sup>	82.6	90.3	96.1	94.3	90.1	87.4	89.9	94.0	94.7	91.3	78.5	89.9
3D MT network <sup>[15]</sup>	87.1	93.2	97.0	94.6	90.2	88.3	91.1	93.8	96.5	96.0	85.7	92.1
Gaitset <sup>[16]</sup>	90.8	97.9	99.4	96.9	93.6	91.7	95.0	97.8	98.9	96.8	85.8	95.0
J-CNN <sup>[17]</sup>	87.2	93.2	96.3	95.9	91.6	86.5	89.8	93.8	95.1	93.0	80.8	91.2
ours	<b>97.0</b>	<b>99.0</b>	<b>100</b>	<b>98.0</b>	<b>98.0</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>99.0</b>	<b>95.0</b>	<b>98.5</b>
Probe BG #1-2	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	Mean
LB-subGEI <sup>[15]</sup>	64.2	80.6	82.7	76.9	64.8	63.1	68.0	76.9	82.2	75.4	61.3	72.4
GaitSet <sup>[16]</sup>	83.8	91.2	<b>91.8</b>	<b>88.8</b>	83.3	81.0	<b>84.1</b>	<b>90.0</b>	<b>92.2</b>	<b>94.4</b>	79.0	87.2
J-CNN <sup>[17]</sup>	73.1	78.1	83.8	81.6	71.6	65.5	71.0	80.7	79.1	78.6	68.0	75.0
ours	<b>89.0</b>	<b>93.9</b>	88.9	85.7	<b>93.0</b>	<b>90.0</b>	83.0	85.0	87.0	86.8	<b>81.0</b>	<b>87.5</b>
Probe CL #1-2	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	Mean
LB-subGEI <sup>[15]</sup>	37.7	57.2	66.6	61.1	55.2	54.6	55.2	59.1	58.9	48.8	39.4	53.98
GaitSet <sup>[16]</sup>	<b>61.4</b>	<b>75.4</b>	<b>80.7</b>	<b>77.3</b>	72.1	<b>70.1</b>	<b>71.5</b>	<b>73.5</b>	<b>73.5</b>	<b>68.4</b>	<b>50.0</b>	<b>70.4</b>
J-CNN <sup>[17]</sup>	46.1	58.4	64.4	64.2	55.5	50.5	54.7	55.8	53.3	51.3	39.9	54.01
ours	52.0	51.0	55.0	59.0	<b>73.0</b>	69.0	66.0	58.0	49.0	50.0	41.0	56.6

从表 5 可以看出,在正常行走(NM)与携带背包(BG)的条件下,本文的方法在识别效果上有一定提升,说明本文方法在无遮挡或少量遮挡条件下能提取更有鉴别性的步态特征;在穿着外套(CL)的条件下,与同为以类能量图作为输入特征的算法 LB-subGEI<sup>[15]</sup>和 J-CNN<sup>[17]</sup>相比,识别效果有所提升,但是与以视频序列作为输入特征的算法 GaitSet<sup>[16]</sup>相比,只在 72°、90°视角有较高的识别率,这是因为大衣(CL)对于人体遮挡范围较大,有效的步态特征难以提取,需要较多的样本才能进行区分,而本文的方法以 GEI 作为输入,相比以视频序列作为输入,数据规模缩小了几十倍,因此导致了识别效果较差。

为了在单一变量条件下验证本文方法对携带背包与穿着大衣的影响,对去除视角的变化进行了实验,以测试对象的同视角的 nm01 ~ nm04 作为 gallery 进行实验,结果如图 4 所示。

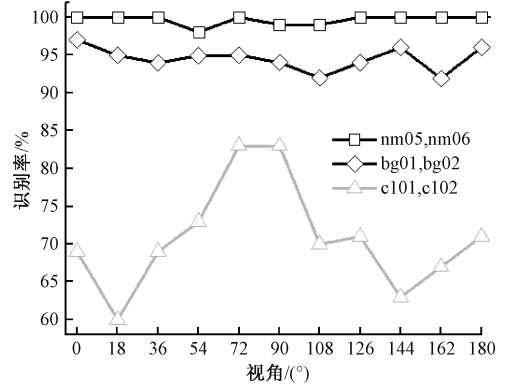


图 4 在 CASIA-B 的同视角下的 rank-1 准确率

Fig. 4 Accuracy of rank-1 under the same perspective of CASIA-B

图 4 中,在 NM、BG、CL 条件下准确率的均值分别为

99.7%、94.5%、70.8%，可以明显看出本文方法在携带背包的条件下仍有很高的识别率，而对于衣着条件下72°与90°视角的条件下识别率较高，其余的识别率比较低，原因可能是72°与90°视角的步态特征更为完善，也可能是数据集中关于衣着的训练数据相对较少。为了进一步验证本文方法的有效性和潜在应用，对本文的方法在CASIA-B数据集上采用等错误率(equal error rate, EER)进行评估，得到了EER为1.4。并且对全部数据进行了测试，测试时间大约为2 min，批处理速度27 fps，已经能够满足大部分步态识别系统的实时监测需求。

### 2.3 在OU-ISIR Treadmill B的实验与分析

OU-ISIR Treadmill B是以服装组合为主的数据集，其数据均是90°视角的步态，因此本文在该数据集的28种服装组合上对所提出的方法进行验证，结果如图5所示。

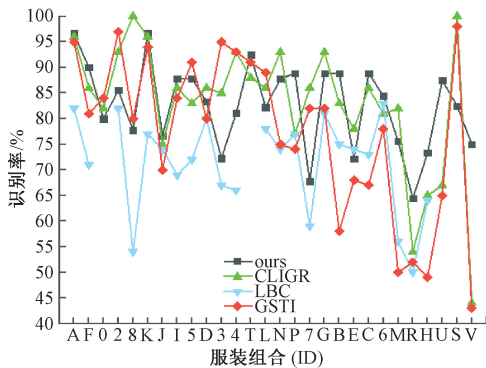


图5 在OU-ISIR Treadmill B上不同服装组合的rank-1准确率

Fig. 5 The rank-1 accuracy rate of different clothing combinations on OU-ISIR Treadmill B

从图5可以看出，本文的方法在众多服装组合影响下，相比当前比较先进的算法GSTI<sup>[18]</sup>、LBC<sup>[19]</sup>与CLIGR<sup>[20]</sup>，仍然有较高的识别率，并且识别效果相比其他几个更加平稳，在最难识别的几种服装组合上，如R(雨衣)、H(休闲裤+羽绒服)、U(短裙+风衣)、V(短裙+羽绒服)的情况下，本文的方法相比其他最好的识别效果分别提升了10%、8%、20%、31%。但是在3(常规裤+半衫+帽子)、4(常规裤+半衫+鸭舌帽)、7(常规裤+长大衣+帽子)的情况下，识别率低于其他的算法，通过对训练数据进行分析，发现在总共50个对象的881个视频序列中，这3种分别只有8、8、9个视频序列，而少量的输入在训练时对参数变化的影响较小，导致了识别率的下降。其EER为4.98，验证了本文方法在应对服装鲁棒性的鉴别能力，在表6中显示了本文与当前识别率较高的算法的平均准确率的对比。

表6 在OU-ISIR Treadmill B上与其他方法的比较  
Table 6 Comparison with other methods on the OU-ISIR Gait Database, the Treadmill Dataset B

Algorithm	Rank-1/%
GSTI <sup>[18]</sup>	77.25
LBC <sup>[19]</sup>	74.4
CLIGR <sup>[20]</sup>	82.13
ours	82.6

### 3 结论

本文利用残差网络构建了一个用于克服携带背包与衣着变化的步态识别学习框架。通过将多级分块结构、Inception模块引入残差网络模型，使得模型能够从GEI图像的输入中提取有鉴别性的特征。使用GEI作为残差网络的原始输入，在公开数据集CASIA-B和OU-ISIR Treadmill B上进行了实验验证。实验结果表明，本文的方法能有效提取72°与90°视角的步态特征，在无遮挡的情况下，对特征进行水平多级分块能有效提取步态特征，可以达到98.5%的识别率；对于携带背包的小区域遮挡，通过分块可以保证在训练时当前层的其他区域的卷积核不受遮挡区域的影响，并减弱遮挡区域对于下一层卷积的影响，所以对于小区域遮挡仍有较高的识别率；对于大范围的遮挡，通过分块可以保证在训练时当前层的无遮挡的区域特征能被有效提取，但由于被遮挡区域过大，在进入下一层卷积运算时受到的影响较大，所以提升效果较小。因此，本文的方法在实际应用中有一定的借鉴意义。在未来的工作中，将进一步改进网络结构，研究更有效的提取具有服装不变性的特征。

### 参考文献

- [1] 周倩,孙运强,姚爱琴,等. 基于人体关节定位的步态识别技术研究[J]. 国外电子测量技术, 2019, 38(1): 52-56.  
ZHOU Q, SUN Y Q, YAO AI Q, et al. Research on gait recognition technology based on human joint point positioning [J]. Foreign Electronic Measurement Technology, 2019, 38(1): 52-56.
- [2] LYNNERUP N, LARSEN P K. Gait as evidence [J]. IET Biometrics, 2014, 3(2): 47-54.
- [3] 贾晓辉,王涛,刘今越,等. 基于人体模型映射的步态识别及意图感知方法[J]. 仪器仪表学报, 2020, 41(12): 236-244.  
JIA X H, WANG T, LIU J Y, et al. Gait recognition and intention perception method based on human model mapping [J]. Chinese Journal of Scientific Instrument, 2020, 41(12): 236-244.
- [4] ISLAM M S, ISLAM M R, HOSSAIN M A, et al.

- Subband entropy-based features for clothing invariant human gait recognition[J]. *Advanced Robotics*, 2017, 31(10): 519-530.
- [5] THAPAR D, JASWAL G, NIGAM A, et al. Gait metric learning siamese network exploiting dual of spatio-temporal 3D-CNN intra and LSTM based inter gait-cycle-segment features[J]. *Pattern Recognition Letters*, 2019, 125: 646-653.
- [6] HAWAS A R, EL-KHOBBY H A, ABD-ELNABY M, et al. Gait identification by convolutional neural networks and optical flow[J]. *Multimedia Tools and Applications*, 2019, 78(18): 25873-25888.
- [7] LAM T H W, CHEUNG K H, LIU J N K. Gait flow image: A silhouette-based gait representation for human identification[J]. *Pattern Recognition*, 2011, 44(4): 973-987.
- [8] WANG C, ZHANG J P, WANG L, et al. Human identification using temporal information preserving gait template[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11): 2164-2176.
- [9] AL-TAYYAN A, ASSALEH K, SHANABLEH T. Decision-level fusion for single-view gait recognition with various carrying and clothing conditions[J]. *Image and Vision Computing*, 2017, 61: 54-69.
- [10] 卢来, 邓文, 吴卫祖. 基于改进深度卷积神经网络的步态识别算法[J]. *电子测量与仪器学报*, 2019, 33(2): 88-93.  
LU L, DENG W, WU W Z. Gait recognition algorithm based on improved deep convolutional neural network[J]. *Journal of Electronic Measurement and Instrumentation*, 2019, 33(2): 88-93.
- [11] 王杰. 基于卷积神经网络的步态识别算法研究[D]. 西安: 西安科技大学, 2019.  
WANG J. Research on gait recognition algorithm based on convolutional neural network [D]. Xi'an: Xi'an University of Science and Technology, 2019.
- [12] WANG Y N, YANG Y M, LI Y. Recognition and difference analysis of human walking gaits based on intelligent processing of video images[J]. *Traitement du Signal*, 2020, 37(6): 1085-1091.
- [13] 刘太阳. 基于改进残差网络的步态识别研究[D]. 哈尔滨: 哈尔滨理工大学, 2021.  
LIU T Y. Research on gait recognition based on improved residual network [D]. Harbin: Harbin University of Science and Technology, 2021.
- [14] 胡少晖, 王修晖, 刘砚秋. 基于多支路残差深度网络的跨视角步态识别方法[J]. *模式识别与人工智能*, 2021, 34(5): 455-462.  
HU SH H, WANG X H, LIU Y Q. A cross-view gait recognition method based on multi-branch residual deep network [J]. *Pattern Recognition and Artificial Intelligence*, 2021, 34(5): 455-462.
- [15] WU Z F, HUANG Y Z, WANG L, et al. A comprehensive study on cross-view gait based human identification with deep CNNs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39(2): 209-226.
- [16] CHAO H P, HE Y W, ZHANG J P, et al. GaitSet: Regarding gait as a set for cross-view gait recognition[J]. *AAAI Computer Vision and Pattern Recognition*, 2018, 33: 8126-8133.
- [17] ZHANG Y Q, HUANG Y Z, WANG L, et al. A comprehensive study on gait biometrics using a joint CNN-based method[J]. *Pattern Recognition*, 2019, 93: 228-236.
- [18] GHAEMINIA M H, SHOKOUHI S B. On the selection of spatiotemporal filtering with classifier ensemble method for effective gait recognition[J]. *Signal Image & Video Processing*, 2019, 13: 43-51.
- [19] XU Z P, LU W, ZHANG Q, et al. Gait recognition based on capsule network [J]. *Journal of Visual Communication and Image Representation*, 2019, 59: 159-167.
- [20] GHEBLEH A, MOGHADDAM M E. Clothing-invariant human gait recognition using an adaptive outlier detection method [J]. *Journal of Multimedia Tools and Applications*, 2018, 77(7): 8237-8257.

## 作者简介



张红颖(通信作者), 分别于2001年、2004年、2007年在天津大学获得学士、硕士、博士学位, 现在为中国民航大学教授, 主要研究方向为图像工程与计算机视觉。

E-mail: carole\_zhang0716@163.com

**Zhang Hongying** (Corresponding author)

received her B. Sc., M. Sc., and Ph. D. degrees from Tianjin University in 2001, 2004, and 2007 respectively. Now she is a professor at the Civil Aviation University of China. Her main research interests include image engineering and computer vision.



田鹏华, 2019年于青海大学获得学士学位, 现在为中国民航大学研究生, 主要研究方向为步态识别。

E-mail: 371679794@qq.com

**Tian Penghua** received his B. Sc. degree

from Qinghai University in 2019. Now he is a M. Sc. candidate in Civil Aviation University of China. His main research interest includes gait recognition.