

视频监控场景下基于单视角步态的人体身份及属性识别系统^①



廖嘉城, 梁 艳, 王冰冰, 潘家辉

(华南师范大学 软件学院, 佛山 528225)

通讯作者: 潘家辉, E-mail: panjh82@qq.com

摘 要: 基于步态的特征识别是一种新兴的生物特征识别技术, 旨在通过人们走路的姿态进行身份和相关属性的分析. 与其他的生物识别技术相比, 基于步态的识别方法具有难隐藏性、非接触性和可远距离使用的优点. 本文设计出一个视频监控场景下基于单视角步态的人体身份及属性识别系统, 该系统通过图像处理方法从复杂的监控视频中实时检测出人体的步态, 经过利用深度学习训练过的算法进行分析后, 获取人体的身份、性别和年龄信息. 实验表明, 系统的身份识别准确率达 98.1%, 性别预测准确率达 97.1%, 年龄预测平均绝对误差为 6.21 岁, 实验结果均优于传统基准算法, 且系统开发成本低, 支持实时检测, 能充分满足中小规模步态研究与分析的需要.

关键词: 步态; 身份识别; 属性识别; 实时检测; 深度学习

引用格式: 廖嘉城, 梁艳, 王冰冰, 潘家辉. 视频监控场景下基于单视角步态的人体身份及属性识别系统. 计算机系统应用, 2020, 29(8): 113-120. <http://www.c-s-a.org.cn/1003-3254/7571.html>

Single-View Gait-Based Identity and Attributes Recognition System under Video Surveillance

LIAO Jia-Cheng, LIANG Yan, WANG Bing-Bing, PAN Jia-Hui

(School of Software, South China Normal University, Foshan 528225, China)

Abstract: Gait-based feature recognition is an emerging biometric authentication technology, aiming at analyzing human characteristics such as identity through the walking posture of people. Compared with other biological recognition technologies, gait-based methods have the advantages of being difficult to hide, contactless, and remotely usable. This study designs a single-view gait-based human identity and attributes recognition system under video surveillance. The system uses image processing methods to detect a human gait in real-time from a complex surveillance video. After analyzing with the algorithm trained by deep learning, it can obtain the information of human's identity, gender, and age. Experiments show that the accuracy rate of the system is 98.1%, the accuracy of gender prediction is 97.1%, and the mean absolute error of the age prediction is 6.21, which are better than the traditional benchmark. The system is costless, supporting real-time detection, which can fully meet the needs of small and medium-scale gait research and analysis.

Key words: gait; identity recognition; attributes recognition; real-time detection; deep learning

1 引言

1.1 研究背景

随着信息技术的不断发展, 人们在日常生活中需要

更频繁地利用特定的身份识别技术验证人的身份. 近几年, 基于人脸、指纹、虹膜等生物特征的生物识别方法日益成熟, 且由于其便捷可靠的特点, 越来越多的生活

^① 基金项目: 国家自然科学基金面上项目 (61876067); 广东省自然科学基金面上项目 (2019A1515011375); 广东省科技创新人才专项珠江科技新星专题 (201710010038)

Foundation item: General Program of National Natural Science Foundation of China (61876067); General Program of Natural Science Foundation of Guangdong Province (2019A1515011375); Special Fund for S&T Innovative Talent under Pearl River S&T Star of Guangdong Province (201710010038)

收稿时间: 2020-01-28; 修改时间: 2020-02-27; 采用时间: 2020-03-11; csa 在线出版时间: 2020-07-29

场景也利用此类技术去完成身份识别的任务.但与此同时,上述几类技术都需要被识别者与特定设备在短距离间实现交互,当有人刻意去避免近距离接触时,以上的方法将无法进行识别甚至无法反馈被识别者有价值的信息.面对此类局限,步态识别(gait recognition)因具备难隐藏性、非接触性和可远距离使用等特点,逐渐成为生物识别领域中一项可行的识别方案^[1,2].

步态识别是一项能利用人的身体体型和步行姿态识别身份和预测人的性别、年龄等相关属性的技术.相关研究表明,由于人们在各项生理条件存在细微的差异,每个人走路姿势及其走路“风格”也会截然不同,因此想伪装他人的走路姿势并不容易^[3].近年来,随着视频监控设备的普及,步态识别在社会安全、市场营销、生物认证、视频监控和法律援助等领域逐渐发挥更重要的作用.

1.2 国内外研究现状

目前,步态识别技术的实现主要分为两大类.一类是以人工特征建模为基础的传统机器学习算法,另一类则是基于深度学习算法.

在传统的步态识别研究中,能否从人的步态视频中提取到具有明显区分性的步态特征,将直接影响到最终实验结果的好坏.一部分研究者^[4,5]重点关注人行走时人体局部的变化,以更为简单的轨迹计算代替步态特征.Han等^[6]和Lu等^[7]则将多个步态帧转换为类能量图的方式去描述反映步态特征.Makihara等^[8]通过构建视角转化模型将步态图像转化为视角独立、对象独立的特征.提取到步态特征后,研究者会通过支持向量机(Support Vector Machine, SVM)等分类方法对特征进行相应的区分,最终构建特征与被识别者身份、性别以及年龄等属性的映射关系.由于步态识别研究早期并没有较大规模的开源数据集,因此上述的传统机器学习算法都是不同步态数据集上进行的实验.虽然这些方法在各自的数据集中基于步态的身份识别任务都能达到90%上的准确率,但很难将他们统一进行优劣的比较.

随着时间的推移,人们在研究过程中逐渐发现,传统方法会受到人体外在因素如视角、环境条件、服装等干扰,研究者并无法做到对每一种特殊情况进行相应的建模.与此同时,随着大数据技术的兴起以及硬件条件的不断升级,一方面,许多步态识别研究队

伍开源了相关的数据集(例如中科院的CASIA-B大规模多视角步态数据集^[9]),算法研究有了统一且权威的评价指标;另一方面,深度学习在图像/视频等领域取得了非凡的突破,越来越多研究者开始利用深度学习提升步态识别算法的效果以及健壮性.Wu等^[10]围绕步态能量图(Gait Energy Image, GEI)^[6]和卷积神经网络(Convolutional Neural Networks, CNN),充分利用CNN强大的自适应与自学习能力直接训练算法模型,但由于网络结构简单且当时缺乏有效训练神经网络的方法,该方法整体识别率一般,在CASIA-B上的90°步态的识别率为81.5%,效果仅比传统方法好一些.Battistone等^[11]选择步态周期内连续的帧作为训练输入,利用长短期记忆网络(Long Short-Term Memory, LSTM)挖掘步态帧之间内部联系,提出TGLSTM网络结构,在CASIA-B上的平均识别率为86.4%,但该方法以多个帧作为输入需要较大的计算量,无法满足工程应用实时性需求.Zhang等^[12]利用迁移学习,在VGG-D^[13]模型的基础上进行微调,最后在CASIA-B上的平均识别率为95.7%.

在应用领域方面,美国MiniSun公司研发出的IDEEA生活步态系统可以在自然工作和生活状态下记录使用者的多种步态参数(单脚支撑时间、步长、速度等),通过不同指标较高精度地测量出人体健康程度.国内的银河水滴科技可实现远距离多角度的步态识别,还能完成超大范围人群密度测算与实时计数,广泛应用于安防、公共交通、商业等场景.但上述系统存在一定的局限性:一方面上述系统开发成本高,对于相关需求要求较低的小团队、小企业来说,难以支付其昂贵的费用;另一方面,上述系统应用专注于人体身份的识别,在身份识别失效的情况下无法返回反馈信息提供使用者参考.因此,本文设计了一款成本较低、支持实时检测、可通过步态准确识别出人物身份以及性别、年龄等相关属性的系统,它可灵活根据实际需要进行功能调整与修改,能很好满足一些小团队、小企业的实际使用和开发需要,具有重大的现实意义.

2 系统设计

2.1 系统结构

本系统可分成4个模块(如图1):预处理、特征提取、算法模型训练以及系统实现.

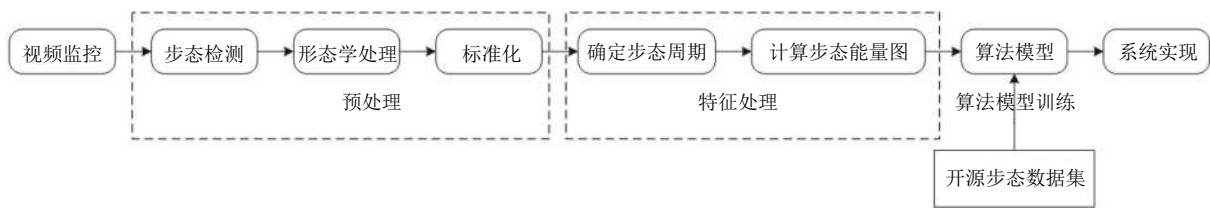


图1 系统设计结构图

在预处理环节中, 我们将从人行走的步态视频中分离前景与背景, 把提取出来的前景经过形态学处理后形成完整的步态周期, 并利用 Han 等^[6]的方法将步态周期形成步态能量图, 并以此作为算法模型训练和实际应用的特征进行身份和属性的识别; 算法模型训练是使用中科院 CASIA-B 大规模多视角步态数据集^[9]和大阪大学 OU-ISIR 大型步态数据集^[14], 利用卷积神经网络训练相应的算法模型, 在实际系统中, 需要用到基于步态的身份识别、性别识别、年龄识别 3 种算法模型以完成系统所需要实现的功能; 在系统实现中, 我们提供一个 Windows 系统下的客户端, 用户可以对录制好的步态视频进行步态的分析, 也可以连接摄像头, 实时地从监控画面进行步态的分析。

2.2 预处理

在预处理过程中, 我们需要将连续的步态帧转化为单一的步态能量图。第 1 步是步态检测, 即是从序列图像中将人体步行区域从背景图像中提取出来。考虑到项目的实际情况, 在此系统中, 本文采用背景减除法提取人步行前景。背景减除法^[15]是一种有效的运动对象检测算法, 基本思想是利用背景的参数模型来近似背景图像的像素值, 将当前帧与背景图像进行差分比较实现对运动区域的检测, 其中区别较大的像素区域被认为是运动区域, 而区别较小的像素区域被认为是背景区域。

在动态的视频中由于受环境的影响, 一般算法检测到的步态轮廓大多存在偏差, 可能出现噪声、孔洞等现象。为获得更清晰、边缘部分更平滑的步态轮廓, 本文利用形态学中的腐蚀和膨胀操作对二值化的轮廓图像进行处理: 首先利用腐蚀操作去除图像中一些较小的噪声点, 然后通过膨胀操作填充图像的孔洞, 从而有效地提高图像质量^[16]。

虽然此时我们已经能够将步态帧转化为清晰的人体步行前景, 但行人轮廓仅占整幅图像很小的一部分,

冗余的背景会极大地影响算法模型训练以及系统分析的效率。针对该问题, 本文先从整幅图像中提取出行人轮廓, 接着将轮廓放置于图像的中心位置并充斥图像的大部分区域, 最后把整幅图的大小标准化为 128×88 像素。

2.3 特征提取

在进行预处理后, 需要将一个步态周期内的所有帧转换为步态能量图。步态能量图表示步态序列在时间与空间标准化形成的二维图像, 与步态序列相比, 步态能量图不仅在一定程度上保留了时间信息, 大大减少步态的数据量, 而且对单个步态帧中的噪声也不敏感^[6]。步态能量图 $G(x, y)$ 可依据式 (1) 获得。

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

其中, x 和 y 表示二维图像的坐标值, N 表示一个步态周期内步态帧的个数, $B_t(x, y)$ 表示单个步态帧。本文中所有图像大小都统一标准化为 128×88 像素。

用归一化自相关 (Normalized Auto Correlation, NAC) 函数的方法可以匹配不同步态帧的相似程度, 进而确定一个完整的步态周期^[17]。但此方法需使用特定的公式, 计算量也颇大。考虑人行走的步态姿势人手臂与腿的运动具有一定的规律性, 本文采取计算步态帧中的人体轮廓“宽高比”的方式进行步态周期的估算。某个人体轮廓的“宽高比” r_i^k 可依据式 (2) 获得。

$$r_i^k = \frac{w_i^k}{h_i^k} \quad (2)$$

其中, k 表示该行人步态序列中步态周期的个数, i 表示第 k 个步态周期中的步态帧数, w_i^k 和 h_i^k 表示第 k 个步态周期中第 i 帧人体轮廓的宽和高。若 T_k 表示第 k 个步骤周期的所有步态帧, 则

$$T^k = r_1^k, r_2^k, \dots, r_n^k \quad (3)$$

其中, n 表示一个步态周期内步态帧的总数, 即便同一

个个体,不同的步态周期中,步态帧总数不一定相等.在一个步态周期中,人体轮廓“宽高比”需要达到两次极大值和极小值:

$$r_i^k = r_{\max}^k, r_j^k = r_{\max}^k, i \neq j \quad (4)$$

$$r_i^k = r_{\min}^k, r_j^k = r_{\min}^k, i \neq j \quad (5)$$

其中, r_{\max}^k 表示步态周期内人体轮廓“宽高比”的极大值,此时人的手臂以及腿张开角度达到最大; r_{\min}^k 表示步态周期内人体轮廓“宽高比”的极小值,此时人呈站立的状态.与此同时,步态周期第一帧与最后一帧的差值需小于等于阈值 α (本系统取 $\alpha=0.03$), 即:

$$|r_1^k - r_n^k| \leq \alpha \quad (6)$$

2.4 算法模型训练

本系统的算法训练统一使用 CNN 作为核心架构. CNN 是一种适用于大规模图像及视频识别、理解、分类等计算机视觉领域的一种经典且高效的深度学习网络结构.

具体网络架构方面,本系统参考文献 [18] 提出的

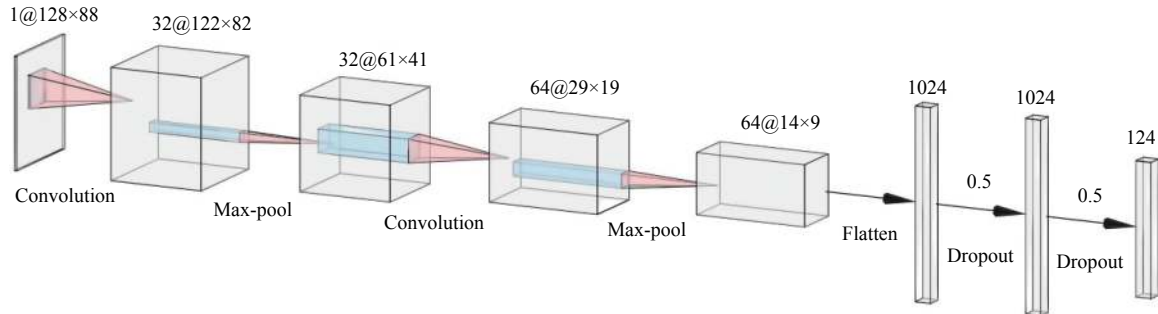


图2 本系统采用的基础 CNN 结构

2.5 系统实现

本系统按照“客户端+服务器+数据库”的思路进行设计与构建.在 Windows 系统上,项目通过 PyQt 进行客户端界面的绘制,客户端上的代码以 Python 实现.在客户端上,系统主要采集已有步态视频或者是实时采集到的监控视频中的步态信息,在转换为步态能量图后,传送给服务器.在本地环境中,加载了算法模型的服务器会挂起等待着客户端传送的数据,待有数据从端口传送进来的时候,经过运算,返回步态图像所对应的行人的性别、年龄等信息.客户端还会将从服务器端返回的特征信息与私有步态数据库进行相似度比

网络结构,并在他们的基础上对网络模型进行了调整并修改了相关参数,提升了身份识别的准确率,并用相同的网络结构进行性别/年龄预测模型的训练.本系统所用于算法训练的基本神经网络结构如图 2 所示:神经网络的输入是标准化后的步态能量图 (128×88 像素);随后利用两组“卷积层+池化层”提取步态的特征,前一组卷积层使用 32 个 7×7 的卷积核,步长为 1,池化层为 2×2,步长为 2,后一层卷积层使用 64 个 5×5 的卷积核,步长为 2,池化层为 3×3,步长为 2;利用 Flatten 层转化为一维数据后,利用两层拥有 1024 个神经元的全连接层进行信息的整合与特征区分,最后进行分类与回归.其中,在身份识别和性别预测算法模型训练中,我们根据数据集的人数划分成相应数目的类,从而将识别人身份的任务转化为分类的任务;而在年龄预测算法模型训练中,我们利用此网络完成回归任务.分类任务使用交叉熵作为损失函数,回归任务使用平均绝对误差作为损失函数,训练使用 Adam 优化器, batch-size 为 128,训练轮次以及使用的激活函数随着任务的不同而变化.

对,进一步确认行人的身份.最后客户端上会显示出行人的身份、性别以及年龄信息.具体实现效果将在下一章详细展示.

3 实验结果

3.1 运行环境

本系统在 Windows 64 位操作系统上成功运行与测试,算法部分通过 Python、C++ 编译实现,并在 Colaboratory 平台上进行训练,系统界面利用 Qt Creator 完成开发;在硬件层面,本系统运行以及测试的电脑 CPU 为 i5-8300H,内存 8 GB,显卡为 1050Ti,本系统所用的

摄像头是 1080P USB 摄像头。

3.2 步态数据集

3.2.1 CASIA-B 大规模多视角步态数据集

本系统使用中科院 CASIA-B 大规模多视角步态数据集训练身份识别算法模型。数据集中采集人数为 124 人, 93 名男性, 31 名女性, 每人从 0° 到 180° 分 11 个视角, 每个视角间隔 18° 的方式采集。每个人行走状态又分为 3 种: 穿大衣、携带包裹、正常条件。每个视频分辨率为 320 像素 \times 240 像素, 25 帧/s。由于本项目只是对 90° 下的步态识别进行研究, 因此算法模型仅使用数据集中 90° 下的数据进行训练。

3.2.2 OU-ISIR 大型步态数据集

本系统使用大阪大学 (OU) 科学与工业研究所 (ISIR) 的大型步态数据集训练基于步态的性别与年龄预测算法模型。该数据集的数据包含年龄、性别标签, 适用于评估与年龄、性别相关的人体步态分析算法 (例如, 以步态为基础的年龄预测的一般领域的研究工作估计)。该数据集由 63 846 名受试者组成, 他们沿着摄像机拍摄的路线行走, 行走视频 30 帧/s, 分辨率 640 像素 \times 480 像素, 受试者年龄范围为 2 至 90 岁。在数据集中, 每个受试者行走序列都转化为一个尺寸标准化的步态能量图。

3.3 步态检测效果

本文通过背景减除法以及 2.3 的方法, 从既有或者实时的视频中获取人体步态能量图。但采用不同的背景减除算法去除背景的效果以及效率不同, 选用合适的背景减除算法能帮助系统准确且实时地计算步态能量图, 提升身份及属性识别的准确率。常用的背景减除算法有以下 3 种: 基于高斯混合模型背景分割算法 (Gaussian mixture-based background segmentation algorithm, MOG)^[19]、基于高斯混合模型背景分割改进算法 (MOG2)^[20] 以及基于 K 邻近的背景分割算法 (K-Nearest Neighbours-based background segmentation algorithm, KNN)^[21]。

图 3 展示的是 3 种不同的算法在参数相同的情况下对同一段视频去除背景后使用 2.3 节中的方法计算得到的步态能量图, 表 1 展示了 3 种算法结合 2.3 节中的方法获取同一段视频中行人步态能量图的时间。

进行测试的源视频时长约为 5 s, 人在行走过程中, 视频背景会产生镜像进行干扰, 因此可很好地测试不

同背景减除算法的健壮性。结合图 3 和表 1 可看出, 使用 KNN 方法最后获取到的步态能量图更为清晰和完整。虽然 MOG2 算法最后计算时间更短, 但使用该算法最后生成的步态能量图效果很糟糕, 难以作为后续识别的输入。在 5 s 的视频中, KNN 结合 2.3 中方法仅用 1.66 s 即可生成对应的步态能量图, 意味着在一段正常的行人步行时间内, 系统有足够的时间实时地捕获到人的步态能量图以进行后续的工作。

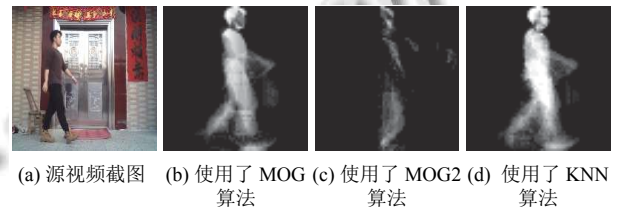


图 3 3 种背景减除算法配合 2.3 节中方法得到的步态能量图

表 1 不同算法结合 2.3 节中方法计算步态能量图表

算法	时间(s)
MOG	2.269
MOG2	1.538
KNN	1.663

3.4 算法识别效果

本文将步态数据集中连续的步态帧转换为步态能量图, 并作为 CNN 的输入训练算法模型。由于训练的样本较少, 我们将训练集、验证集、测试集按照 8:1:1 进行划分, 以尽可能提供较多的数据进行训练。3 种算法模型在 Colaboratory 平台上进行训练的时间均在一个小时之内, 在测试的主机上进行训练的时间最长也不超过两个小时。因此在硬件资源受限的情况下, 本文提出的算法依然能够被快速训练并运用, 能够满足中小团队的开发需要。

基于步态的身份识别算法模型训练过程如图 4 所示, 图中“acc”表示算法在训练集上的识别准确率, “val_acc”表示算法在验证集上的识别准确率。由图可知, 经过 25 轮训练之后, 在验证集的识别准确率可达到 97.7%, 算法识别准确率高, 收敛速度快, 且不会出现过拟合现象。由表 2 可知, 本文提出的基于步态的身份识别方法最终在测试集的识别率达到 98.1%, 优于 Battistone 等^[11] 和 Zhang 等^[12] 的方法, 说明能达到较好的实际应用效果。

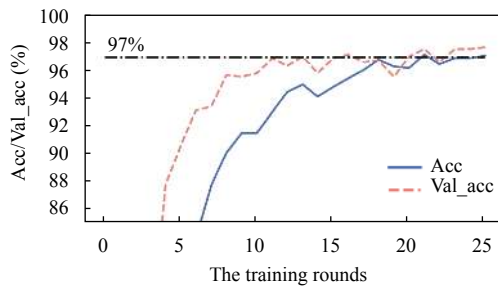


图4 基于步态的身份识别算法模型训练过程图

表2 基于步态的身份识别算法识别率对比

方法	识别率(%)
Battistone等 ^[11]	86.4
Zhang等 ^[12]	95.7
本文方法	98.1

性别预测算法模型训练过程如图5所示。由于用于性别识别和年龄预测的数据集比较大,因此一轮的训练后验证集识别准确率可达92.6%,且基于步态的性别预测本质上是二分类问题,分类任务较为简单,最后算法整体识别率高,经过25轮训练之后,识别率可达到97%左右,轻微过拟合。因为目前尚未有文献对该任务进行专门的研究,因此该算法并没有一个可比较参考的基准,但从实验效果来看,该算法模型最终在测试集的识别率达到97.1%,说明对于年龄的预测具有极高的准确性和鲁棒性。

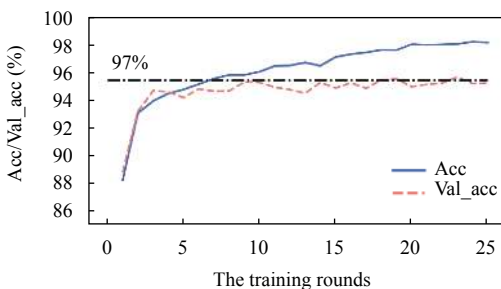


图5 基于步态的性别预测算法模型训练过程图

年龄预测算法模型训练过程如图6所示。图中“mae”表示算法在训练集上的平均绝对误差,“val_mae”表示算法在验证集上的平均绝对误差。由图可知,在经过18轮的训练后,训练集上的平均绝对误差逐渐下降且依然有不断下降的趋势,但在验证集的平均绝对误差并没随着训练轮次的增加有明显的下降趋势,说明对于验证集来说,算法已经接近收敛,过拟合现象明显。

我们将本文的方法与传统的两种基准算法进行对比,实验结果参考表3。其中,“GPR”表示使用高斯过程回归(Gaussian Process Regression)^[22]算法计算的平均绝对误差,“SVR”表示使用支持向量回归(Support Vector Regression)^[23]算法计算的平均绝对误差,其中“GPR”和“SVR”是Xu等^[14]提出的两种基于步态的年龄预测基准算法。由表可知,本文提出的方法,在测试集的平均绝对误差为6.21岁,效果较好,明显优于传统方法。

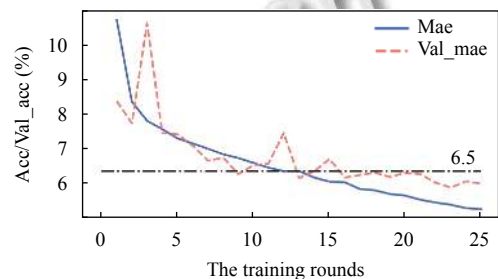


图6 基于步态的年龄预测算法模型训练过程图

表3 基于步态的年龄预测结果对比

方法	平均绝对误差(岁)
GPR ^[22]	7.30
SVR ^[23]	7.66
本文方法	6.21

步态能量图最大程度地保留了人体行走的空间和时间信息,而卷积神经网络可通过卷积以及非线性运算挖掘步态能量图中深层特征,最终显著地区分出不同步态的属性信息。在本节实验中,利用步态能量图与卷积神经网络进行身份、性别以及年龄的识别/预测算法模型的训练。算法经过测试集的测试,身份识别的准确率达98.1%,性别预测准确率达97.1%,年龄预测平均绝对误差在6.21岁,说明本文提出的方法对于单视角行人身份及属性的分析具备较强的鲁棒性和容错能力。

3.5 在线系统测试

本系统将在4种不同的场景中进行相对应的测试,图7~图9这3幅图是系统对已经录制好的步态视频进行测试的效果,其中,图7中的测试背景与另外两种情景不同。图10是本系统通过外置摄像头进行实时检测行人并识别行人身份和属性信息的测试效果。不同场景都各有一人参与测试,受试者为3名男性,1名女性,年龄皆在20~23岁之间。



图7 对步态视频进行识别(身份已知且性别为男)



图8 对步态视频进行识别(身份已知且性别为女)



图9 对步态视频进行识别(身份未知)



图10 实时环境下系统利用外接摄像头进行步态识别

从测试结果可以看出,在不同的背景条件下,本系统能够在上述4种不同的状态下从既有或者实时的视频中准确地检测出行人的位置并进行定位与跟踪,说明系统具备对于背景变化的抗干扰能力.对已在本地数据库中登记的受试者,本系统能正确地识别其身份和性别;对未在本数据库登记的受试者,本系统也能反馈其性别和年龄信息.此外,通过外置摄像头,本系统能实时捕捉行人的步态信息,并进行相应识别与预测.由于算法模型是训练好加载到本地服务器中,因

此当客户端输入相应的步态能量图,服务器端可快速反馈相对应的步态分析.上述所有的测试场景,系统在2~2.5 s内可分析获得相对应人体的步态信息,因此可以满足实际场景中实时性检测的需要.虽然在年龄预测环节,本系统在部分场景中预测行人的年龄与真实年龄的偏差较大,如图8和图9所示的场景,预测的年龄误差约有5岁左右.但总体上看,本系统开发成本较低,可支持人体步态的实时检测,算法具备较好的鲁棒性,能准确反馈行人信息,可满足部分场景的使用需求,具有实际开发意义.

4 讨论与总结

本文构建和开发了一款可在视频监控场景下通过步态准确识别出人物身份以及反馈相关特征信息的系统.该系统运用卷积神经网络进行训练,在Shiraga^[18]的基础上进行改进和优化,在测试集上,身份识别的准确率达98.1%,性别预测准确率达97.1%,年龄预测平均绝对误差在6.21岁,基本达到实际场景中对于步态识别与分析的要求.与此同时,该系统开发成本较低,支持实时检测,可根据实际需要对其功能进行灵活地调整与修改,因此能很好满足一些小团队、小企业的实际使用和开发需要,具有重要的现实意义.

虽然本系统已达到较为理想的效果,但依然有许多地方可以进一步改进和提升:一方面,在更为复杂的监控环境中,背景、光线等外部环境变换难以估计,行人的行走方向也并不会因为固定好摄像机的角度而保持不变,所以步态识别的前期预处理本身就是庞大且复杂的工程,需要更具鲁棒性的方法进行预处理并提取到步态分析所用到的步态能量图;其次,更加精准的步态识别与分析,也需要更加庞大的数据集,以及更好的模型训练方法,令算法可以在精度及实时性等方面满足更复杂场景的需要.

参考文献

- 1 胡海明, 翟晓梅. 生物识别技术应用的伦理问题研究综述. 科学与社会, 2018, 8(3): 49-59, 68.
- 2 Wan CS, Wang L, Phoha VV. A survey on gait recognition. ACM Computing Surveys, 2018, 51(5): 89.
- 3 张诚. 基于深度学习的步态识别关键技术研究[硕士学位论文]. 北京: 北京邮电大学, 2016.
- 4 Jean F, Albu AB, Bergevin R. Towards view-invariant gait modeling: Computing view-normalized body part

- trajectories. *Pattern Recognition*, 2009, 42(11): 2936–2949. [doi: [10.1016/j.patcog.2009.05.006](https://doi.org/10.1016/j.patcog.2009.05.006)]
- 5 Ng H, Tan WH, Abdullah J, *et al.* Development of vision based multiview gait recognition system with MMUGait database. *The Scientific World Journal*, 2014, 2014: 376569.
- 6 Han J, Bhanu B. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(2): 316–322. [doi: [10.1109/TPAMI.2006.38](https://doi.org/10.1109/TPAMI.2006.38)]
- 7 Lu JW, Wang G, Moulin P. Human identity and gender recognition from gait sequences with arbitrary walking directions. *IEEE Transactions on Information Forensics and Security*, 2014, 9(1): 51–61. [doi: [10.1109/TIFS.2013.2291969](https://doi.org/10.1109/TIFS.2013.2291969)]
- 8 Makihara Y, Sagawa R, Mukaigawa Y, *et al.* Gait recognition using a view transformation model in the frequency domain. *Proceedings of the 9th European Conference on Computer Vision*. Graz, Austria. 2006. 151–163.
- 9 Yu SQ, Tan DL, Tan TN. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. *Proceedings of the 18th International Conference on Pattern Recognition*. Hong Kong, China. 2006. 441–444.
- 10 Wu ZF, Huang YZ, Wang L. Learning representative deep features for image set analysis. *IEEE Transactions on Multimedia*, 2015, 17(11): 1960–1968. [doi: [10.1109/TMM.2015.2477681](https://doi.org/10.1109/TMM.2015.2477681)]
- 11 Battistone F, Petrosino A. TGLSTM: A time based graph deep learning approach to gait recognition. *Pattern Recognition Letters*, 2019, 126: 132–138. [doi: [10.1016/j.patrec.2018.05.004](https://doi.org/10.1016/j.patrec.2018.05.004)]
- 12 Zhang XF, Sun SQ, Li C, *et al.* DeepGait: A learning deep convolutional representation for gait recognition. *Proceedings of the 12th Chinese Conference on Biometric Recognition*. Shenzhen, China. 2017. 447–456.
- 13 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv: 1409.1556*, 2014.
- 14 Xu C, Makihara Y, Ogi G, *et al.* The OU-ISIR gait database comprising the large population dataset with age and performance evaluation of age estimation. *IPSJ Transactions on Computer Vision and Applications*, 2017, 9(1): 24. [doi: [10.1186/s41074-017-0035-2](https://doi.org/10.1186/s41074-017-0035-2)]
- 15 张洁. 智能视频监控中的背景减除算法研究 [硕士学位论文]. 成都: 电子科技大学, 2015.
- 16 支双双, 赵庆会, 唐璘. 多层语义融合 CNN 的步态人体语义分割. *中国图象图形学报*, 2019, 24(8): 1302–1314. [doi: [10.11834/jig.180597](https://doi.org/10.11834/jig.180597)]
- 17 李超. 基于卷积神经网络的人体行为分析与步态识别研究 [博士学位论文]. 杭州: 浙江大学, 2019.
- 18 Shiraga K, Makihara Y, Muramatsu D, *et al.* GEINet: View-invariant gait recognition using a convolutional neural network; *Proceedings of 2016 International Conference on Biometrics*. Halmstad, Sweden. 2016.1–8.
- 19 KaewTrakulpong P, Bowden R. An improved adaptive background mixture model for real-time tracking with shadow detection. In: Remagnino P, Jones GA, Paragios N, *et al.* eds. *Video-Based Surveillance Systems*. Boston, MA: Springer. 2002. 135–144.
- 20 Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction. *Proceedings of the 17th International Conference on Pattern Recognition*. Cambridge, UK. 2004.28–31.
- 21 Zivkovic Z, Van Der Heijden F. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 2006, 27(7): 773–780. [doi: [10.1016/j.patrec.2005.11.005](https://doi.org/10.1016/j.patrec.2005.11.005)]
- 22 Makihara Y, Okumura M, Iwama H, *et al.* Gait-based age estimation using a whole-generation gait database. *Proceedings of 2011 International Joint Conference on Biometrics*. Washington, WA, USA. 2011. 1–6.
- 23 Smola AJ, Schölkopf B. A tutorial on support vector regression. *Statistics and Computing*, 2004, 14(3): 199–222. [doi: [10.1023/B:STCO.0000035301.49549.88](https://doi.org/10.1023/B:STCO.0000035301.49549.88)]