

基于耦合神经网络的步态识别方法

沈 喆, 盛 阳

(沈阳航空航天大学民用航空学院, 沈阳 110136)

摘要: 主流步态识别过程中识别率易受到视角变化的影响,使用耦合网络模型解决最小类间距离大于最大类内距离的问题,利用步态样本比对解决识别问题。用二值图像序列合成步态能量图,联合使用逻辑回归与对比损失函数对模型进行训练优化,实验验证耦合神经网络的步态识别性能,并与普通卷积神经网络模型(CNNs)的识别结果进行对比。在背包与大衣遮挡情况下的识别率达到 73.7%和 60.5%,高于 CNNs 的识别率,并提高了遮挡情况下步态识别精确度。

关键词: 步态识别; 步态能量图; 视角变化; 耦合神经网络模型

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1671-1807(2024)09-0124-04

2022 年 10 月,《信息安全技术 步态识别数据安全要求》(GB/T 41773—2022)正式发布实施。由于生活中某些特定的场所对安全等级有特殊要求,但是由于人员数量比较庞大,且实际情况较为复杂导致了传统的生物特征身份识别很难继续发挥效能^[1],环球网 2019 年报道的“步态识别:50 米外,秒读你的‘活体密码’”,通过人的体型以及走路姿态来识别行人身份的步态识别技术,相比于其他类似人脸、指纹等传统生物特征进行身份识别方法来说其具有无需近身识别和极不易被模仿的优点^[2]。现在普遍使用的步态识别方法按照研究图像类型不同分为基于非模型识别方法与基于模型识别方法^[3-4],其中非模型识别方法直接从人体轮廓中提取步态特征,通过构建步态轮廓图的帧与帧之间的关系针对步态图像的轮廓信息进行分析,利用步态样本轮廓得出步态特征。与模型步态识别法相比较,非模型步态识别方法具有模型简单、整体算力低等优点,对目标图像或者视频的分辨率要求更低,也是如今步态识别主要采用的方法。非模型步态识别法不需要考虑人体内部骨架等结构如何运动,重点研究这些轮廓信息的外部轮廓动态方式,研究步态图像中步态轮廓像素量的变化量,根据变化量得到对应的步态特征来实现步态识别。刘晓阳和靖薇^[5]顺利将步态识别运用于识别煤矿井下人员身份,采用自适应的度量学习实现特征提取,为煤矿

井下人员提供位置自动化检测; Derlatka 和 Borowska^[6]提出异构基分类器集合的方法进行步态识别,并提高了准确率与识别时间。

1 图像预处理

1.1 形态学处理

实验采用 CASIA-B 数据库中经过运动目标提取且二值化后的图像,虽不需要进行行人检测以及背景剔除,但仍存在噪声、粘连等问题,要对这些二值化的图像进行形态学处理。其中膨胀操作可以将图像中的物体变大,使它更加连通;腐蚀操作则可以将图像中的物体变小,使它更加细化;开运算可以做到去除噪点,将图像边缘的锯齿消除;闭运算可以填补图像中物体的孔洞。腐蚀: $A \odot B = \{x, y \mid (B)_{xy} \subseteq A\}$, 膨胀: $A \oplus B = \{x, y \mid (B)_{xy} \cap A \neq \emptyset\}$, 分别表示用一个 B 结构来腐蚀 A 图像和用一个 B 结构来膨胀 A 图像。 \odot 和 \oplus 分别表示腐蚀操作和膨胀操作, (x, y) 为锚点的位置, x 和 y 为结构元值为 1 的像素相对锚点的位置偏移。

1.2 图像的归一化

由于步态数据采集过程中设备是固定不动的,采集到的图样是由小到大再到小的状态,图像中也因此存在过多噪声点,而步态归一化^[7]通过成比例调整步态轮廓的高度,把图像中步态轮廓的中心位置和大小调整至一致。通过找到图像中上下左右像素值的最大值点,由此生成一个足以覆盖步态轮廓的最大矩形

收稿日期: 2023-12-31

基金项目: 辽宁省教育厅青年科技人才“育苗”项目(JYT2020130);痕迹检验鉴定技术公安部重点实验室开放课题(HJKF201907);公安部文件检验重点实验室开放课题(FTKF202102)

作者简介: 沈喆(1980—),女,辽宁鞍山人,博士,讲师,研究方向为图像处理;盛阳(1999—),男,辽宁大连人,硕士研究生,研究方向为交通信息工程及控制。

框,通过矩形框裁剪出人体目标的轮廓图像,将其高度放缩至 64 像素并保持此纵横比不变来实现。如式(1)所示计算调整轮廓图的质心标 x_c 。归一化操作也可以优化步态图像中特征受到的外界背景环境以及采样光照的影响,解决步态图像中局部过暗过亮的问题,抑制住了步态图像中的干扰噪点。

$$x_c = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

式中: x_c 为轮廓图质心坐标; x_i 为像素值为 1 的点的横坐标; N 为整幅图中像素值为 1 的点的总数。

2 步态能量图

步态能量图^[8]的原理是使用多通道映射函数来使步态图像的更多时空信息在均值的过程中保留下来,通过选取人的一只脚尖刚离地到另外一只脚尖刚离地这一段时间作为一个步态周期,使用一个周期内步态序列轮廓图求一个平均灰度值图像,根据式(2)得出步态能量图。

$$GEI = \frac{1}{N} \sum_{t=1}^N B_t(x, y) \quad (2)$$

式中: N 为一系列步态轮廓序列的个数; t 为当前时刻的序列图; x, y 为图像坐标。

图 1 为一个步态周期和其步态能量图,图中像素点越亮的部分代表其在步态周期中出现得越频繁。



图 1 一个步态周期以及步态能量图

3 主流 CNNs 方法

普通卷积神经网络模型(CNNs)通常由卷积层、池化层、下采样层、全连接层和输出层组成。将图像作为输入量输入到第 1 个卷积层中,将卷积层的输出作为激活函数。然后通过添加池化层来减少输入参数量,在输出层得到输出值之前,经过的卷积层有助于提取特征。CNNs 中的输出层属于全连接层,来自其他层的输入向量在全连接层被平化,这一步的作用是方便后续输出转换为卷积神经网络需要的参数量。

卷积层对图像各个子区域进行特征提取,使用激活函数对消除高层偏差后输出的特征向量进行转换,卷积层特征提取计算公式为

$$F^i = \text{ReLU}(w^i \otimes F^{i-1} + \beta) \quad (3)$$

式中: F^i 为模型提取输出的特征向量; F^{i-1} 为子区域输出的特征向量; w^i 为偏差权重; β 为偏差向量,线性整流函数(ReLU)作为激活函数,定义为 $\text{ReLU} = \max(0, x)$,其中 x 为神经网络的输入。

这里的池化层的作用是克服卷积时对位置的敏感,通常采用最大值池化或平均值池化。在绝大多数应用情景下,最大值池化提取特征能力较强, $\max P$ 表示最大池操作,操作定义为

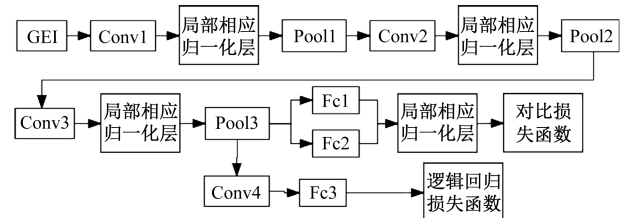
$$F^i = \max P(F^{i-1})$$

式中: F^{i-1} 为池化区域内的输入像素点; F^i 为区域内特征的集合。

4 耦合神经网络模型

耦合神经网络^[9]包括双层卷积神经网络、对比损失层^[10]和逻辑回归损失层,如图 2 所示。图 2 显示,来自卷积层的特性向量,通过全连接层输出至对比损失层,通过识别输入样本类别来训练网络模型参数。其将从上层提取的步态特征向量与 Conv4 区分并连接,对比损失层每次输入一对图片,进入同一个网络进行特征提取。对应到图中就是上下均为一对。那么两个网络就会提取两个特征,然后用这两个网络输出的特征进行相减来提取特征差异,然后再经过全连接层作一个二分类,判断输入图像是否为同一人,最后耦合神经网络通过损失函数组合训练。这里使用损失函数还可以帮助识别模型是否存在过拟合或欠拟合的情况,如果损失函数在训练集上较低但在验证集上较高,可能表示模型存在过拟合。

GEI 作为样本输入到卷积网络单元 Conv1 中,扩张操作后输入到局部归一化层 N 和 Pool1 中。卷积网络单元 Conv2 与 Conv3 对池化层的输出量获取步态特征向量的方式一样,将卷积单元 Conv3 经过池化层 Pool3 得出的特征向量输入到 Fc1 和 Fc2 中,然后输入到局部响应归一化层当中,对比损失层输入得到的特征向量。接下来,通过对来自 Pool3 输出层的两个特征图进行差分融合,将卷积网络的 Pool3 输出层的特征向量分别输入到 Conv4 和对比损失层,将其映射为 0~1 分布,确定步态样本输入的相似度。



GEI 为步态能量图;Conv1、Conv2、Conv3 为卷积网络单元; Pool1、Pool2、Pool3 为池化层;Fc1、Fc2、Fc3 为全连接层

图 2 耦合神经网络流程

5 网络训练方法

结合逻辑回归^[11]和对比损失函数^[12]对网络进行参数训练,训练后提取样本步态特征来进一步实现对比、识别行人身份。经过全连接层归一化后的特征向量输入到对比损失层中,其中对比损失函数的定义如式(4)所示, α 为手动设置的阈值,用来计算网络中两个输出向量之间的度量距离。

$$L_c = yd_{I_1, I_2}^2 + (1 - y)(\alpha - d_{I_1, I_2}^2)_+^2 + (z)_+; \max(z, 0) \quad (4)$$

式中: I_1, I_2 为样本对的差值特征向量; L_c 为通过网络特征提取预测到的差值特征向量的概率分布; y 为两个样本是否匹配的标签; d 为两个样本特征之间的欧氏距离; z 为设定的阈值。

在卷积网络单元 Conv4 中输入得到的特征向量进行差值融合,把 Fc3 输出的特征向量输入到逻辑回归损失函数中,并基于 0~1 分布实现对输入步态样本的身份识别。逻辑回归函数常用于二分类,损失函数越小,模型在训练集上面的拟合效果越好,逻辑回归损失函数如公式(5)所示,其中 θ 的最优值可以通过梯度下降的方法,得到的 θ 在训练集中的拟合效果非常好,但是可能会存在过拟合的现象,这时可以使用 L_1, L_2 正则化的方法来降低过拟合带来的影响。

$$L(\theta) = -\frac{1}{n} \sum_{i=1}^n [f_i \ln(p_\theta) + (1 - f_i) \ln(1 - p_\theta)] \quad (5)$$

式中: n 为输入样本数量; i 为目标类; p_θ 为损失层参数; f_i 为真实目标样本差值特征向量的概率分布。

耦合神经网络的参数由损失函数关联法驱动,并通过一个特定的参数 λ 将两者结合起来,从而控制损失函数在总损失函数中的实现状态,二者联合如式(6)所示。

$$\text{Loss} = L(\theta) + \lambda L_c \quad (6)$$

λ 取值范围为 $(0, +\infty)$,在模型训练过程中,当 $\lambda = 0$ 时,损失函数的检验与网络参数的匹配无关;当 $\lambda \neq 0$ 时,验证损失函数生效,逻辑回归损失函数与网络参数拟合的比例逐渐减小。可以看到,当 $\lambda \rightarrow +\infty$ 时,可以看作对比损失函数单独涉及对整体参数的调整。经过实验发现当 $\lambda = 0.06$ 时,该参数下的损失函数组合优化的性能最优,步态识别率也得到提高。

6 实验分析

实验运用 MATLAB 2017-a 仿真工具,操作系统为 Windows 10,设备内存容量 16GB, CPU 为酷睿 i5-1340p,采用中科院 CASIA-B 步态数据集作为

训练集。CASIA-B 数据集是使用最广泛的步态数据集,包含 124 个受试者的步态轮廓数据,每个受试者有 10 组视频。每个受试者的步态序列样本分别采集自 $0^\circ, 18^\circ, \dots, 180^\circ$ 的 11 个角度,每个视角下都有 10 个步态序列。选择在 4 种正常步态条件下收集的步态序列作为训练样本,记为 NM1-4。选择了两个行走序列作为验证样本,记为 NM5-6。验证样本集选择大衣遮挡下的步态序列(CL1-2)与背包遮挡下的步态序列(BG1-2),通过参数 λ 将两个损失函数联合训练网络参数。图 3 为耦合网络模型对正常行走、背包遮挡以及大衣遮挡下的步态识别率。通过图 4 可以看出,在背包情况下耦合网络对各视角的识别综合情况要好于 CNNs。

对 CNNs^[13]的识别数据与通过本文使用的耦合神经网络得出的结果进行综合对比,见表 1。

由表 1 可以看出,耦合神经网络的平均识别率要高于 CNNs,在考虑 11 个不同视角下的 NM、CL、BG 3 种不同的行走情况时发现,在某些行走情况和

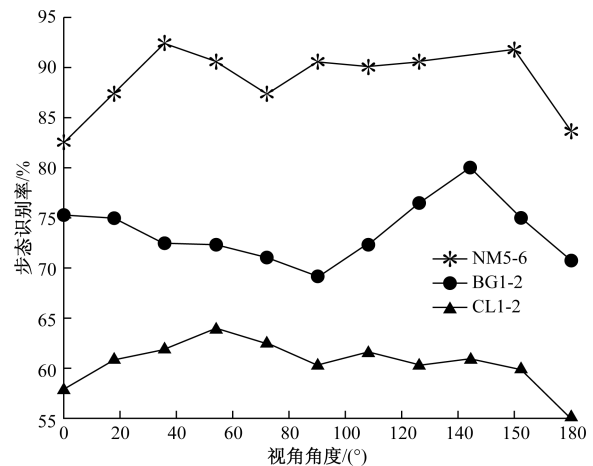


图3 耦合在各情况下的识别率

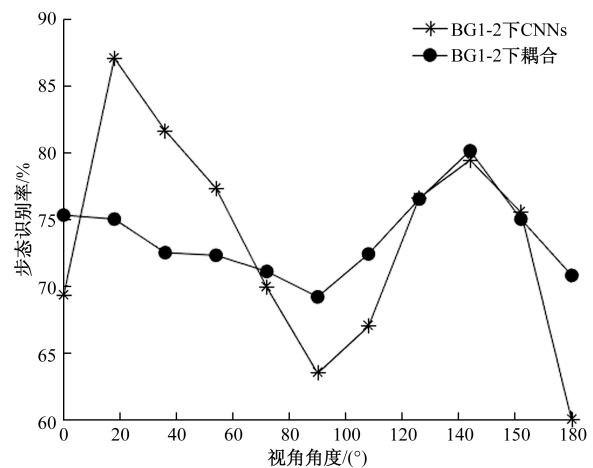


图4 背包情况下CNNs与耦合网络识别率对比

表1 不同视角下不同条件步态识别率

探针 NM1-4		识别率/%											平均值/%
		0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
NM5-6	CNNs	79.2	88.3	95.5	92.6	88.9	86.9	89.3	92.7	92.2	89.3	75.3	88.2
	耦合	82.4	87.3	92.4	90.5	87.2	90.5	90.0	90.5	91.1	90.4	83.6	88.7
BG1-2	CNNs	69.3	87.0	81.6	77.3	69.9	63.5	67.0	76.6	79.4	75.5	60.0	73.3
	耦合	75.3	75.0	72.5	72.3	71.1	69.2	72.4	76.5	80.1	75.0	70.8	73.7
CL1-2	CNNs	37.9	53.6	60.0	58.4	55.2	52.0	52.9	53.5	53.0	48.1	37.5	51.1
	耦合	57.9	61.0	62.0	64	62.6	60.4	61.6	60.4	61.0	60.0	55.0	60.5

视角,如探针为 NM、BG 条件下的 $18^\circ\sim 72^\circ$ 时,耦合网络识别率低于 CNNs。CNNs 对该一段视角的识别更加稳定,但是耦合网络模型综合性能优于 CNNs,证明了模型的有效性。通过损失函数联合训练参数可以有效解决步态识别中最小的类间距离大于最大的类内距离的问题,有效提高了在不同视角下的步态识别率。

7 结论

提出基于耦合网络模型的步态识别方法,介绍了步态识别的原理以及耦合神经网络的基本架构,对 CASIA-B 数据集所提供的步态图像进行图像归一化等处理并由此得出步态能量图,将得到的步态能量图输入到构建的耦合神经网络模型中提取步态特征差异,构建对比损失函数和逻辑回归损失函数,并经过训练得出 λ ,使得损失函数联合优化的性能达到最优,使用训练后的网络模型来进行步态特征识别比对,能够有效提升步态识别精度。最后,通过实验得出耦合神经网络在验证样本集 NM5-6、BG1-2 以及 CL1-2 上的步态识别率,并与传统 CNNs 进行实验结果对比,验证了步态识别时当角度发生变化或人物出现外物遮挡时,耦合神经网络模型可以有效地减少步态识别效果的误差。

参考文献

[1] 赵红柏. 智慧安防小区管控系统方案设计与实现[D]. 杭州: 浙江工业大学, 2021.

- [2] 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.
- [3] SINGH J P, JAIN S, ARORA S, et al. A survey of behavioral biometric gait recognition: current success and future perspectives[J]. Archives of Computational Methods in Engineering, 2019, 28(1): 1-42.
- [4] 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.
- [5] 刘晓阳, 靖薇. 基于步态识别的煤矿井下人员身份识别方法; 201810936703.3[P]. 2024-01-22.
- [6] DERLATKA M, BOROWSKA M. Ensemble of heterogeneous base classifiers for human gait recognition[J]. Sensors (Basel, Switzerland), 2023, 23(1): 2-4.
- [7] 赵鑫泽, 代雪晶. 基于轻量级卷积网络的步态识别研究[J]. 现代计算机, 2023, 29(13): 60-64.
- [8] 刘亮亮. 基于深度学习和步态能量图的步态识别研究[D]. 哈尔滨: 哈尔滨工程大学, 2022.
- [9] 童随兵. 基于深度网络的步态识别技术研究[D]. 上海: 上海交通大学, 2020.
- [10] 赵启林, 卓家寿. 非线性方程组的耦合神经网络算法[J]. 河海大学学报(自然科学版), 2000(5): 38-40.
- [11] 王琳琳, 梁凤梅, 刘阿建. 基于卷积神经网络特征重加权的行人再辨识[J]. 小型微型计算机系统, 2019, 40(4): 834-838.
- [12] 陈曙. 基于特征转换学习的卡口车牌检测识别方法[D]. 西安: 西安电子科技大学, 2020.
- [13] WU Z, HUANG Y, WANG L, et al. A comprehensive study on cross-view gait based human identification with deep CNNs[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(2): 209-226.

Gait Recognition Method Based on Coupled Neural Network

SHEN Zhe, SHENG Yang

(College of Civil Aviation, Shenyang Aerospace University, Shenyang 110136, China)

Abstract: In order to solve the problem that the recognition rate is easily affected by the change of view Angle in the mainstream gait recognition process, the coupled network model is used to solve the problem that the minimum distance between classes is greater than the maximum distance within classes, and the comparison of gait samples is used to solve the recognition problem. The gait energy diagram is synthesized by binary image sequences, and the model is trained and optimized by logistic regression and contrast loss function. The gait recognition performance of the coupled neural network was verified by experiments and compared with the recognition results of the common convolutional neural network (CNNs). The recognition rate reached 73.7% and 60.5% in the case of backpack and overcoat occlusion, which is higher than CNNs, and the accuracy of gait recognition under occlusion is improved.

Keywords: gait recognition; gait energy diagram; change of perspective; coupled neural network model