



(12) 发明专利

(10) 授权公告号 CN 109255339 B

(45) 授权公告日 2021.04.06

(21) 申请号 201811222012.3

G06N 3/00 (2006.01)

(22) 申请日 2018.10.19

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 109255339 A

- WO 0139655 A2, 2001.06.07
- US 9811720 B2, 2017.11.07
- EP 2439492 A1, 2012.04.11
- CN 104200200 A, 2014.12.10
- CN 105574510 A, 2016.05.11
- CN 101241546 A, 2008.08.13
- CN 107212890 A, 2017.09.29
- CN 104299012 A, 2015.01.21
- CN 103473539 A, 2013.12.25

(43) 申请公布日 2019.01.22

(73) 专利权人 西安电子科技大学
地址 710071 陕西省西安市雁塔区太白南路2号

Zhi-Hua Zhou 等. Deep Forest: Towards An Alternative to Deep Neural Networks. 《arXiv:1702.08835v2》. arXiv, 2017, 1-10.

(72) 发明人 赵盼盼 盛立杰 苗启广 马悦 庞博 秦丹 陈红颖 徐劲夫

(74) 专利代理机构 陕西电子工业专利中心
61205
代理人 田文英 王品华

审查员 林浩

(51) Int. Cl.

G06K 9/00 (2006.01)

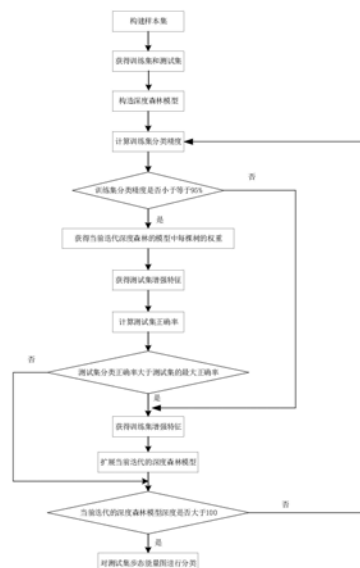
权利要求书4页 说明书8页 附图1页

(54) 发明名称

基于自适应深度森林人体步态能量图分类方法

(57) 摘要

本发明公开了一种基于自适应深度森林人体步态能量图分类方法,其步骤为:(1)构建训练集和测试集;(2)构造自适应深度森林模型;(3)计算自适应深度森林模型分类精度;(4)判断自适应深度森林模型是否达到要求;(5)若训练集准确率及深度未达到要求,扩展自适应深度森林模型,求解自定义二次凸优化问题得到自适应深度森林模型在训练集的增强特征,通过训练集上的增强特征得到扩展后自适应深度森林的训练集;(6)计算测试集分类正确率。本发明减少了模型的复杂度和模型需要训练的权重,同时减小了训练模型所需要的步态能量图样本规模。



1. 一种基于自适应深度森林人体步态能量图的分类方法,其特征在于,求解自适应深度森林模型中每棵树的权重,获取自适应深度森林模型的增强特征,根据深度森林模型的深度自适应的确定深度森林模型复杂度,利用获得深度森林模型的增强特征训练深度森林模型;该方法的具体步骤包括如下:

(1) 构建样本集:

(1a) 对输入的124个目标人物视频进行背景建模与背景实时更新,得到每个目标人物的人体轮廓前景图像;

(1b) 从每个目标人物视频的人体轮廓前景图像中提取人体步态图;

(1c) 分别将每个步态周期内得到的每个目标人物的所有人体轮廓图像的灰度值相加后,再除以每个目标人物的人体轮廓图像的总数,得到与该步态周期对应的步态能量图;

(1d) 取每个目标人物在正常行走条件下且视角为90度的六张步态能量图,共744张步态能量图构成样本集;

(2) 获得训练集和测试集:

(2a) 利用随机组合抽样拼接法,对样本集进行随机采样,得到992个正样本和负样本,将所有的正样本和负样本组成训练集;

(2b) 利用随机组合抽样拼接法,对样本集进行随机采样,得到992个正样本和负样本,将所有的正样本和负样本组成测试集;

(3) 构造自适应深度森林模型:

(3a) 搭建一个由两个随机森林和两个完全随机森林组成的自适应深度森林模型;

(3b) 将自适应深度森林模型中的两个随机森林和两个完全随机森林中树的个数均设置为400棵、树的最大深度均设置为8,自举样标志均设置为True,并行处理的核数均设置为服务器的最大内核数,将自适应深度森林模型的深度设置为h;

(4) 计算当前迭代的自适应深度森林模型分类精度:

(4a) 将当前迭代的自适应深度森林模型训练集输入到当前迭代的自适应深度森林模型中,输出训练集中每个样本对应的类概率;

(4b) 利用训练集分类精度计算公式,计算训练集经过当前迭代的自适应深度森林模型训练后深度森林模型分类精度;

(5) 判断当前迭代的自适应深度森林模型分类精度是否小于95%;若是,执行步骤(6),否则,执行步骤(12);

(6) 获得当前迭代的自适应深度森林模型中每棵树的权重:

(6a) 利用下述的自定义二次凸优化表达式,得到当前迭代的自适应深度森林模型各个随机森林中每棵树的权重:

$$\min_{\xi_i, w^k} \sum_{i=1}^K (\xi_i^k)^2 + \lambda \|w^k\|^2$$

$$\text{s.t. } \xi_i^k \geq \sum_{i=1}^{T_i} P_i^{(t,k)} w^{(t,k)} + \tau$$

$$\xi_i^k \geq 0, i \in K$$

其中, \min_{ξ_i, w^k} 表示求使目标函数式最小化时 ξ_i, w^k 的操作, i 表示训练集中样本的标号, K 表

示训练集样本的总数, Σ 表示求和操作, ξ_i^k 表示训练集中第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林中的基本损失, λ 表示正则化系数, $\|\cdot\|$ 表示 2 范数操作, w^k 表示当前迭代的自适应深度森林模型中第 k 个森林的总权重, s. t. 表示约束条件符号, t 表示当前迭代的自适应深度森林模型中第 k 个随机森林中树的下标, T_k 表示当前迭代的自适应深度森林模型中第 k 个随机森林中树的总数, $P_i^{(t,k)}$ 表示第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中的基本损失, $w^{(t,k)}$ 表示在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 棵树的权重, τ 表示分类间隔;

(6b) 将当前迭代的自适应深度森林模型各个随机森林中每棵树的权重组成当前迭代的自适应深度森林模型的权重;

(7) 获得当前迭代的自适应深度森林模型在测试集上的增强特征:

(7a) 根据当前迭代的自适应深度森林模型中每棵树的权重, 通过当前迭代的自适应深度森林模型的增强特征求解公式, 分别获得测试集中的每个样本为正样本和负样本的增强特征;

(7b) 由测试集中的所有样本的正样本和负样本的增强特征构成当前迭代的自适应深度模型的增强特征;

(8) 计算测试集的正确率:

根据当前迭代的自适应深度森林模型在测试集上的增强特征, 利用当前迭代的自适应深度森林模型测试集正确率计算方法计算测试集的正确率;

(9) 判断当前迭代的自适应深度森林模型测试集正确率是否大于测试集的最大正确率, 若是, 则执行步骤 (10), 否则, 执行步骤 (12);

(10) 获得当前迭代的自适应深度森林模型在训练集上的增强特征:

(10a) 根据当前迭代的自适应深度森林模型中每棵树的权重, 通过当前迭代的自适应深度森林模型的增强特征求解公式, 分别获得训练集中的每个样本为正样本和负样本的增强特征;

(10b) 由训练集中的所有样本的正样本和负样本的增强特征构成当前迭代的自适应深度森林模型的增强特征;

(11) 扩展当前迭代的自适应深度森林模型:

(11a) 将当前迭代的自适应深度森林模型的训练集与当前迭代的自适应深度模型的增强特征首尾拼接, 得到当前迭代的自适应深度森林模型的训练集;

(11b) 将当前迭代的自适应深度森林模型的深度 $h+1$ 作为当前迭代时的自适应深度森林模型的深度 h ;

(12) 判断当前迭代的自适应深度森林模型的深度 h 是否大于 100, 若是, 则得到训练好的深度森林模型后执行步骤 (13), 否则, 执行步骤 (4);

(13) 对测试集中所有的步态能量图进行分类:

将当前迭代的自适应深度森林模型的测试集输入到训练好的深度森林模型中, 对测试集中步态能量图进行分类。

2. 根据权利要求 1 所述的基于自适应深度森林人体步态能量图的分类方法, 其特征在于: 步骤 (1a) 中所述的背景建模与背景实时更新是指, 将人物目标视频的前帧与背景帧相减, 得到目标人物视频的人体轮廓前景图像; 调用视觉库中的功能函数, 将人体轮廓前景图

像中存在的空洞和零散白点去除。

3. 根据权利要求1所述的基于自适应深度森林人体步态能量图的分类方法,其特征在于:步骤(1b)中所述的提取人体步态图的步骤为,将每个人体轮廓最左面的点与最右面的点的距离作为人体轮廓的宽度,最上面的点与最下面的点的距离作为人体轮廓的高度;利用人体的轮廓宽度随时间发生同步周期性改变的特性,通过人体轮廓的宽度变化信号来划分步态周期;按人体轮廓的宽度和高度将人体轮廓图像抠出并保持高宽比不变,将所有人体轮廓图像缩放到同一大小。

4. 根据权利要求1所述的基于自适应深度森林人体步态能量图的分类方法,其特征在于:步骤(2)中所述随机组合抽样拼接法的步骤如下:

第一步,从样本集的124个人物目标中随机选取一个未选过的人物目标的6张步态能量图作为当前处理对象;

第二步,从当前处理对象中,随机选取四对步态能量图,将每对步态能量图进行横向拼接,构成一个正样本并将其标签设置为0,每个目标人物的正样本数量为4;

第三步,从当前处理对象中,随机选取一张步态能量图;

第四步,从样本集中除当前目标人物以外,分别随机选取4个不同目标人物,从所选取的目标人物各自的6张步态能量图中,分别随机选取一张步态能量图,共4张步态能量图;

第五步,用第三步得到的一张步态能量图分别与第四步得到的4张步态能量图两两配对,构成4对步态能量图,将每对步态能量图进行横向拼接,构成一个负样本并将其标签设置为1,每个人物目标的负样本数量为4;

第六步:将当前处理对象标记为已处理;

第七步:对124个目标人物,若每个目标人物的6张步态能量图都标记为已处理,那么共得到496张正样本和496张负样本,即一共992张样本,对124个目标人物,若存在目标人物的6张步态能量图未处理完毕,执行第一步。

5. 根据权利要求1所述的基于自适应深度森林人体步态能量图的分类方法,其特征在于:步骤(6a)中所述第*i*个样本在当前迭代的自适应深度森林模型中第*k*个随机森林的第*t*个树中的基本损失表达式为:

$$P_i^{(t,k)} = Z_i \left(p_{i,0}^{(t,k)} - p_{i,1}^{(t,k)} \right)$$

其中, Z_i 表示指示变量,若训练集中第*i*个样本的标签为0,则 $Z_i = -1$,若训练集中第*i*个样本的标签为1,则 $Z_i = 1$,*表示乘法操作, $p_{i,0}^{(t,k)}$ 表示训练集中第*i*个样本在当前迭代的自适应深度森林模型中第*k*个随机森林的第*t*个树中正样本的概率, $p_{i,1}^{(t,k)}$ 表示训练集中第*i*个样本在当前迭代的自适应深度森林模型中第*k*个随机森林的第*t*个树中负样本的概率。

6. 根据权利要求5所述的基于自适应深度森林人体步态能量图的分类方法,其特征在于:步骤(7a)中所述当前迭代的自适应深度森林模型的增强特征求解公式如下:

$$v_{i,0} = \sum_{t=1}^{T_k} p_{i,0}^{(t,k)} w^{(t,k)}$$

$$v_{i,1} = \sum_{t=1}^{T_k} p_{i,1}^{(t,k)} w^{(t,k)}$$

其中, $v_{i,0}$ 表示训练集中第*i*个样本为正样本的增强概率, $v_{i,1}$ 表示训练集中第*i*个样本

为负样本的增强概率。

7. 根据权利要求1所述的基于自适应深度森林人体步态能量图的分类方法,其特征在于:步骤(8)中所述当前迭代的自适应深度森林模型测试集正确率计算方法如下:

第一步,对测试集中的每个样本,正样本的增强特征数值大于负样本的增强特征值,并且样本的标签为0,则样本分类正确,否则样本分类错误;

第二步,对测试集中的每个样本,正样本的增强特征数值小于负样本的增强特征值,并且样本的标签为1,则样本分类正确,否则样本分类错误;

第三步,统计测试集中被正确分类的样本的总数;

第四步,用测试集中被正确分类的样本的总数除以测试集样本的总数,得到深度森林模型的测试集正确率。

8. 根据权利要求1所述的基于自适应深度森林人体步态能量图的分类方法,其特征在于:步骤(9)中所述测试集的最大正确率是指测试集在深度为 $1, 2, \dots, h$ 的自适应深度森林模型上正确率的最大值。

基于自适应深度森林人体步态能量图的分类方法

技术领域

[0001] 本发明属于图像处理技术领域,更进一步涉及图像识别技术领域中的一种基于自适应深度森林的人体步态能量图分类方法。本发明可用于对人体步态能量图GEI (Gait Energy Image) 中的步态特征进行分类。

背景技术

[0002] 步态识别技术是根据视频序列中人走路的姿态进行身份识别的一种生物技术。由于步态识别具有非侵犯性、远距离识别性以及难以隐藏的特点,因此步态识别在国家公共安全、金融安全、身份认证、视频监控等领域有着广泛的应用前景。

[0003] 银河水滴科技(北京)有限公司在其申请的专利文献“一种基于深度学习的步态识别方法”(专利申请号:201410587758X,申请公开号:CN104299012A)中提出了一种基于深度学习的步态识别方法。该方法采用步态能量图来描述步态序列,通过深度卷积神经网络训练匹配模型,从而匹配步态识别人的身份。该方法的训练过程为:对已标记好身份的涉及多个视角的训练步态视频序列提取步态能量图,重复选取其中任意两个对基于卷积神经网络的匹配模型进行训练直至模型收敛;该方法的识别过程为:对单视角待识别步态视频和已注册步态视频序列分别提取步态能量图,利用训练过程中训练好的基于卷积神经网络的匹配模型计算单视角待识别步态视频的步态能量图与已注册步态视频序列每个步态能量图的相似度,依据相似度的大小进行身份预测,并输出识别结果。该方法对跨较大视角的步态变化有很强的鲁棒性,有效地解决了现有步态识别技术在处理跨视角步态识别时精度不高的问题。但是,该方法仍然存在的不足之处是,由于该方法是基于深度卷积神经网络的分类方法,而基于深度卷积神经网络的分类方法需要大量的训练样本,在训练样本量比较少的情况下,深度卷积神经网络的分类方法不能充分学习人体步态能量图中的步态特征,从而降低了深度卷积神经网络的特征提取能力,这使得深度卷积神经网络的分类方法不适用于小样本的数据,并且基于深度卷积神经网络的分类性能严重依赖于参数调试过程,训练过程中需要消耗较大的计算资源。

[0004] 中国科学院计算技术研究所在其申请的专利文献“步态身份识别方法及系统”(专利申请号:CN201710136803.3,申请公开号:CN107016346A)中提供一种步态识别方法。该方法从由用户携带的智能终端获取的加速度数据中提取行为特征、位置特征和步态特征;利用预先训练好的行为识别模型根据所述行为特征来识别用户当前行为;利用预先训练好的位置识别模型根据所述位置特征和所识别的用户当前行为来识别所述智能终端的当前位置;利用预先训练好的步态识别模型根据所述步态特征、所识别的用户当前行为和所述智能终端的当前位置来识别用户的身份。该方法通过分层递进的识别方式提高了步态身份识别的准确率和鲁棒性,而且不需要对相关传感器摆放的位置与方向进行限制,很灵活且方便使用。但是,该方法仍然存在的不足之处是,由于该方法需要用户携带智能终端来获取加速度数据,利用获取的加速度数据来提取用户的行为特征、位置特征和步态特征,这削弱了步态分类方法的非侵犯性,且该方法由预先训练好的行为识别模型、位置识别模型和步态

识别模型三个模型组成,无法将特征提取全部融入到一个模型当中,无法实现端到端的分类。

发明内容

[0005] 本发明的目的在于针对上述现有技术的不足,提出了一种基于自适应深度森林人体步态能量图的分类方法。本发明与现有其他步态分类方法相比,自适应深度森林模型通过求解每棵树的权重,获得自适应深度森林模型的增强特征,用获得的增强特征训练自适应深度森林模型,使训练所需要的权重较少,训练速度快。在训练过程中根据测试集和训练集的正确率不断增加自适应深度森林的深度,使自适应深度森林的复杂度随着训练过程自适应调节,减少了模型的复杂度,充分利用了计算资源。

[0006] 实现本发明目的的思路是,首先对所有步态视频合成步态能量图,利用随机组合抽样的方法,获得训练集和测试集。然后构建一个初始的自适应深度森林模型,并对模型的参数进行初始化。计算前迭代的自适应深度森林模型的分类精度。判断当前迭代的自适应深度森林模型的在训练集上的分类精度是否满足要求,若满足要求则停止训练,此模型即为所需要的模型,若不满足要求,计算深度森林中每棵树的权重,获得当前迭代的自适应深度森林模型在测试集上的增强特征,若在测试集上的准确率达到要求,那么就停止训练,此模型即为所需要的模型,否则对当前深度森林进行扩展,获得深度森林的训练集增强特征,更新当前深度森林的训练集,增加深度森林的深度,并继续对当前深度森林进行训练,直到获得所需要的模型。

[0007] 本发明的具体步骤如下:

[0008] (1) 构建样本集:

[0009] (1a) 对输入的124个目标人物视频进行背景建模与背景实时更新,得到每个目标人物的人体轮廓前景图像;

[0010] (1b) 从每个目标人物视频的人体轮廓前景图像中提取人体步态图;

[0011] (1c) 分别将每个步态周期内得到的每个目标人物的所有人体轮廓图像的灰度值相加后,再除以每个目标人物的人体轮廓图像的总数,得到与该步态周期对应的步态能量图;

[0012] (1d) 取每个目标人物在正常行走条件下且视角为90度的六张步态能量图,共744张步态能量图构成样本集;

[0013] (2) 获得训练集和测试集:

[0014] (2a) 利用随机组合抽样拼接法,对样本集进行随机采样,得到992个正样本和负样本,将所有的正样本和负样本组成训练集;

[0015] (2b) 利用随机组合抽样拼接法,对样本集进行随机采样,得到992个正样本和负样本,将所有的正样本和负样本组成测试集;

[0016] (3) 构造自适应深度森林模型:

[0017] (3a) 搭建一个由两个随机森林和两个完全随机森林组成的自适应深度森林模型;

[0018] (3b) 将自适应深度森林模型中的两个随机森林和两个完全随机森林中树的个数均设置为400棵、树的最大深度均设置为8,自举标志均设置为True,并行处理的核数均设置为服务器的最大内核数,将自适应深度森林模型的深度设置为h;

[0019] (4) 计算当前迭代的自适应深度森林模型的分类精度：

[0020] (4a) 将当前迭代的自适应深度森林模型的训练集输入到当前迭代的自适应深度森林模型中，输出训练集中每个样本对应的类概率；

[0021] (4b) 利用训练集分类精度计算公式，计算训练集经过当前迭代的自适应深度森林模型训练后自适应深度森林模型的分类精度；

[0022] (5) 判断当前迭代的自适应深度森林模型的分类精度是否大于等于95%；若是，执行步骤(12)，否则，执行步骤(6)；

[0023] (6) 获得当前迭代的自适应深度森林模型中每棵树的权重：

[0024] (6a) 利用下述的自定义二次凸优化表达式，得到当前迭代的自适应深度森林模型各个随机森林中每棵树的权重：

$$[0025] \quad \min_{\xi_i, w^k} \sum_{i=1}^K (\xi_i^k)^2 + \lambda \|w^k\|^2$$

$$[0026] \quad \text{s.t.} \quad \xi_i^k \geq \sum_{t=1}^{T_k} P_i^{(t,k)} w^{(t,k)} + \tau$$

$$[0027] \quad \xi_i^k \geq 0, i \in K$$

[0028] 其中， \min_{ξ_i, w^k} 表示求使目标函数式最小化时 ξ_i, w^k 的操作， i 表示训练集中样本的下标， K 表示训练集样本的总数， Σ 表示求和操作， ξ_i^k 表示训练集中第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林中的基本损失， λ 表示正则化系数， $\|\cdot\|$ 表示2范数操作， w^k 表示当前迭代的自适应深度森林模型中第 k 个森林的总权重， s.t. 表示约束条件符号， t 表示当前迭代的自适应深度森林模型中第 k 个随机森林中树的下标， T_k 表示当前迭代的自适应深度森林模型中第 k 个随机森林中树的总数， $P_i^{(t,k)}$ 表示第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中的基本损失， $w^{(t,k)}$ 表示在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 棵树的权重， τ 表示分类间隔；

[0029] (6b) 将当前迭代的自适应深度森林模型各个随机森林中每棵树的权重组成当前迭代的自适应深度森林模型的权重；

[0030] (7) 获得当前迭代的自适应深度森林模型在测试集上的增强特征：

[0031] (7a) 根据当前迭代的自适应深度森林模型中每棵树的权重，通过当前迭代的自适应深度森林模型的增强特征求解公式，分别获得测试集中的每个样本为正样本和负样本的增强特征；

[0032] (7b) 由测试集中的所有样本的正样本和负样本的增强特征构成当前迭代的自适应深度森林模型的增强特征；

[0033] (8) 计算测试集的正确率：

[0034] 根据当前迭代的自适应深度森林模型在测试集上的增强特征，利用当前迭代的自适应深度森林模型测试集正确率计算方法计算测试集的正确率；

[0035] (9) 判断当前迭代的自适应深度森林模型测试集正确率是否大于测试集的最大正确率，若是，则执行步骤(10)，否则，执行步骤(12)；

[0036] (10) 获得当前迭代的自适应深度森林模型在训练集上的增强特征：

[0037] (10a) 根据当前迭代的自适应深度森林模型中每棵树的权重，通过当前迭代的自

适应深度森林模型的增强特征求解公式,分别获得训练集中的每个样本为正样本和负样本的增强特征;

[0038] (10b)由训练集中的所有样本的正样本和负样本的增强特征构成当前迭代的自适应深度森林模型的增强特征;

[0039] (11)扩展当前迭代的自适应深度森林模型:

[0040] (11a)将当前迭代的自适应深度森林模型的训练集与当前迭代的自适应深度森林模型的增强特征首尾拼接,得到当前迭代的自适应深度森林的训练集;

[0041] (11b)将当前迭代的自适应深度森林模型的深度 $h+1$ 作为当前迭代的自适应深度森林模型的深度 h ;

[0042] (12)判断当前迭代的自适应深度森林模型的深度 h 是否大于100,若是,则得到训练好的自适应深度森林模型后执行步骤(13),否则,执行步骤(4);

[0043] (13)对测试集中所有的步态能量图进行分类:

[0044] 将当前迭代的自适应深度森林模型的测试集输入到训练好的自适应深度森林模型中,对测试集中步态能量图进行分类。

[0045] 本发明与现有技术相比较,具有以下优点:

[0046] 第一,由于本发明通过求解每棵树的权重,获得自适应深度森林模型的增强特征,利用增强特征训练自适应深度森林模型,克服了现有技术在训练神经网络过程中需要训练大量的权重,神经网络模型对小样本的数据集不适用的问题,使得本发明具有训练的权重数量少,训练速度快,适用于小数据集的优点。

[0047] 第二,由于本发明在训练过程中根据测试集和训练集的正确率不断增加自适应深度森林的深度,使自适应深度森林的复杂度随着训练过程自适应调节,克服了现有技术神经网络在训练过程中模型复杂度固定不变的问题,使得本发明减少了模型的复杂度,充分利用了计算资源。

附图说明

[0048] 图1是本发明的流程图。

具体实施方式

[0049] 下面结合附图1,对本发明实现的具体步骤做进一步的描述。

[0050] 步骤1,构建样本集。

[0051] 对输入的124个目标人物视频进行背景建模与背景实时更新,得到每个目标人物的人体轮廓前景图像。

[0052] 所述的背景建模与背景实时更新是指,将人物目标视频的前帧与背景帧相减,得到目标人物视频的人体轮廓前景图像;调用视觉库中的功能函数,将人体轮廓前景图像中存在的空洞和零散白点去除。

[0053] 从每个目标人物视频的人体轮廓前景图像中提取人体步态图。

[0054] 所述的提取人体步态图的步骤为,将每个人体轮廓最左面的点与最右面的点的距离作为人体轮廓的宽度,最上面的点与最下面的点的距离作为人体轮廓的高度;利用人体的轮廓宽度随时间发生同步周期性改变的特性,通过人体轮廓的宽度变化信号来划分步态

周期;按人体轮廓的宽度和高度将人体轮廓图像抠出并保持高宽比不变,将所有人体轮廓图像缩放到同一大小。

[0055] 分别将每个步态周期内得到的每个目标人物的所有人体轮廓图像的灰度值相加后,再除以每个目标人物的人体轮廓图像的总数,得到与该步态周期对应的步态能量图。

[0056] 取每个目标人物在正常行走条件下且视角为90度的六张步态能量图,共744张步态能量图构成样本集。

[0057] 步骤2,获得训练集和测试集。

[0058] 利用随机组合抽样拼接法,对样本集进行随机采样,得到992个正样本和负样本,将所有的正样本和负样本组成训练集。

[0059] 利用随机组合抽样拼接法,对样本集进行随机采样,得到992个正样本和负样本,将所有的正样本和负样本组成测试集。

[0060] 所述随机组合抽样拼接法的步骤如下:

[0061] 第一步,从样本集的124个人物目标中随机选取一个未选过的人物目标的6张步态能量图作为当前处理对象。

[0062] 第二步,从当前处理对象中,随机选取四对步态能量图,将每对步态能量图进行横向拼接,构成一个正样本并将其标签设置为0,每个目标人物的正样本数量为4。

[0063] 第三步,从当前处理对象中,随机选取一张步态能量图。

[0064] 第四步,从样本集中除当前目标人物以外,分别随机选取4个不同目标人物,从所选取的目标人物各自的6张步态能量图中,分别随机选取一张步态能量图,共4张步态能量图。

[0065] 第五步,用第三步得到的一张步态能量图分别与第四步得到的4张态能量图两两配对,构成4对步态能量图,将每对步态能量图进行横向拼接,构成一个负样本并将其标签设置为1,每个人物目标的负样本数量为4。

[0066] 第六步:将当前处理对象标记为已处理。

[0067] 第七步:对124个目标人物,若每个目标人物的6张步态能量图都标记为已处理,那么共得到496张正样本和496张负样本,即一共992张样本,对124个目标人物,若存在目标人物的6张步态能量图未处理完毕,执行第一步。

[0068] 步骤3,构造自适应深度森林模型。

[0069] 搭建一个由两个随机森林和两个完全随机森林组成的自适应深度森林模型。

[0070] 将自适应深度森林模型中的两个随机森林和两个完全随机森林中树的个数均设置为400棵、树的最大深度均设置为8,自举样标志均设置为True,并行处理的核数均设置为服务器的最大内核数,将自适应深度森林模型的深度设置为t。

[0071] 步骤4,计算当前迭代的自适应深度森林模型的分类精度。

[0072] 将当前迭代的自适应深度森林模型的训练集输入到当前迭代的自适应深度森林模型中,输出训练集中每个样本对应的类概率。

[0073] 利用训练集分类精度计算公式,计算训练集经过当前迭代的自适应深度森林模型训练后自适应深度森林模型的分类精度。

[0074] 步骤5,判断当前迭代的自适应深度森林模型的分类精度是否小于95%;若是,执行步骤6,否则,执行步骤12。

[0075] 步骤6,获得当前迭代的自适应深度森林模型中每棵树的权重。

[0076] (6.1)利用下述的自定义二次凸优化表达式,得到当前迭代的自适应深度森林模型各个随机森林中每棵树的权重:

$$[0077] \quad \min_{\xi_i, w^k} \sum_{i=1}^K (\xi_i^k)^2 + \lambda \|w^k\|^2$$

$$[0078] \quad \text{s.t.} \quad \xi_i^k \geq \sum_{t=1}^{T_k} P_i^{(t,k)} w^{(t,k)} + \tau$$

$$[0079] \quad \xi_i^k \geq 0, i \in K$$

[0080] 其中, \min_{ξ_i, w^k} 表示求使目标函数式最小化时 ξ_i, w^k 的操作, i 表示训练集中样本的标号, K 表示训练集样本的总数, Σ 表示求和操作, ξ_i^k 表示训练集中第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林中的基本损失, λ 表示正则化系数, $\|\cdot\|$ 表示2范数操作, w^k 表示当前迭代的自适应深度森林模型中第 k 个森林的总权重, s.t. 表示约束条件符号, t 表示当前迭代的自适应深度森林模型中第 k 个随机森林中树的下标, T_k 表示当前迭代的自适应深度森林模型中第 k 个随机森林中树的总数, $P_i^{(t,k)}$ 表示第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中的基本损失, $w^{(t,k)}$ 表示在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 棵树的权重, τ 表示分类间隔;

[0081] 所述第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中的基本损失表达式为:

$$[0082] \quad P_i^{(t,k)} = Z_i^* (p_{i,0}^{(t,k)} - p_{i,1}^{(t,k)})$$

[0083] 其中, Z_i 表示指示变量,若训练集中第 i 个样本的标签为0,则 $Z_i = -1$,若训练集中第 i 个样本的标签为1,则 $Z_i = 1$, $*$ 表示乘法操作, $p_{i,0}^{(t,k)}$ 表示训练集中第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中正样本的概率, $p_{i,1}^{(t,k)}$ 表示训练集中第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中负样本的概率。

[0084] (6.2)将当前迭代的自适应深度森林模型各个随机森林中每棵树的权重组成当前迭代的自适应深度森林模型的权重;

[0085] 步骤7,获得当前迭代的自适应深度森林模型在测试集上的增强特征。

[0086] 根据当前迭代的自适应深度森林模型中每棵树的权重,通过当前迭代的自适应深度森林模型的增强特征求解公式,分别获得测试集中的每个样本为正样本和负样本的增强特征。

[0087] 由测试集中的所有样本的正样本和负样本的增强特征构成当前迭代的自适应深度森林模型的增强特征。

[0088] 所述当前迭代的自适应深度森林模型的增强特征求解公式如下:

$$[0089] \quad v_{i,0} = \sum_{t=1}^{T_k} p_{i,0}^{(t,k)} w^{(t,k)}$$

$$[0090] \quad v_{i,1} = \sum_{t=1}^{T_k} p_{i,1}^{(t,k)} w^{(t,k)}$$

[0091] 其中, $v_{i,0}$ 表示训练集中第 i 个样本为正样本的增强概率, $p_{i,0}^{(t,k)}$ 表示训练集中第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中正样本的概率, $v_{i,1}$ 表示训练集中第 i 个样本为负样本的增强概率, $p_{i,1}^{(t,k)}$ 表示训练集中第 i 个样本在当前迭代的自适应深度森林模型中第 k 个随机森林的第 t 个树中负样本的概率。

[0092] 步骤8, 计算测试集的正确率。

[0093] 根据当前迭代的自适应深度森林模型在测试集上的增强特征, 利用当前迭代的自适应深度森林模型测试集正确率计算方法计算测试集的正确率。

[0094] 所述当前迭代的自适应深度森林模型测试集正确率计算方法如下:

[0095] 第一步, 对测试集中的每个样本, 正样本的增强特征数值大于负样本的增强特征值, 并且样本的标签为0, 则样本分类正确, 否则样本分类错误。

[0096] 第二步, 对测试集中的每个样本, 正样本的增强特征数值小于负样本的增强特征值, 并且样本的标签为1, 则样本分类正确, 否则样本分类错误。

[0097] 第三步, 统计测试集中被正确分类的样本的总数。

[0098] 第四步, 用测试集中被正确分类的样本的总数除以测试集样本的总数, 得到自适应深度森林模型的测试集正确率。

[0099] 步骤9, 判断当前迭代的自适应深度森林模型测试集正确率是否大于测试集的最大正确率, 若是, 则执行步骤10, 否则, 执行步骤12。

[0100] 所述测试集的最大正确率是指测试集在深度为1, 2, ..., h 的自适应深度森林模型上正确率的最大值。

[0101] 步骤10, 获得当前迭代的自适应深度森林模型在训练集上的增强特征。

[0102] 根据当前迭代的自适应深度森林模型中每棵树的权重, 通过当前迭代的自适应深度森林模型的增强特征求解公式, 分别获得训练集中的每个样本为正样本和负样本的增强特征。

[0103] 由训练集中的所有样本的正样本和负样本的增强特征构成当前迭代的自适应深度森林模型的增强特征。

[0104] 步骤11, 扩展当前迭代的自适应深度森林模型。

[0105] 将当前迭代的自适应深度森林模型的训练集与当前迭代的自适应深度森林模型的增强特征首尾拼接, 得到当前迭代的自适应深度森林模型的训练集。

[0106] 将当前迭代的自适应深度森林模型的深度 $t+1$ 作为当前迭代时的自适应深度森林模型的深度 t 。

[0107] 步骤12, 判断当前迭代的自适应深度森林模型的深度 t 是否大于100, 若是, 则得到训练好的自适应深度森林模型后执行步骤13, 否则, 执行步骤4。

[0108] 步骤13, 对测试集中所有的步态能量图进行分类。

[0109] 将当前迭代的自适应深度森林模型的测试集输入到训练好的自适应深度森林模型中, 对测试集中步态能量图进行分类。

[0110] 下面结合仿真实验对本发明的效果做进一步的说明:

[0111] 1、仿真实验条件:

[0112] 本发明的仿真实验的硬件平台为: Intel Core i7-8700K@3.70GHz CPU、64GB RAM、NVIDIA Geforce GTX1080 Ti GPU。

[0113] 本发明的仿真实验的软件平台为:python2.7和TensorFlow1.4。

[0114] 本发明的仿真实验所用数据为公开的步态识别数据集Dataset B,包括124个目标人物在三种行走条件(普通条件,穿大衣,携带包裹条件)和11个视角下人体的步态的轮廓剪影图,本方法合成124个目标人物的步态能量图,并选取每个目标人物在正常行走且视角为90度的条件下的六张步态能量图,共744张步态能量图构成样本集,用随机组合抽样拼接法,对样本集进行随机采样,构成训练集和测试集。

[0115] 2、仿真内容与结果分析:

[0116] 在本发明的仿真实验中,首先依照本发明的方法合成Dataset B数据库中的每个目标人物的在不同条件下的步态能量图,并选取每个目标人物在正常行走条件下且视角为90度的六张步态能量图,共744张步态能量图构成样本集,用随机组合抽样拼接法,对样本集进行随机采样,组成训练集和测试集。使用训练集训练自适应深度森林模型,然后用使用测试样本在训练好的自适应深度森林模型上进行测试,得到本发明的测试集分类精度。

[0117] 本发明的仿真实验是在相同的数据集下,应用本发明方法和现有技术随机森林(Leo Breiman的论文“Random Forests”2001年《Mach.Learning》)对本发明构建的测试集中的步态能量图进行分类。

[0118] 为了验证本发明的方法和现有技术随机森林的分类效果,利用下式,分别计算本发明和随机森林的分类精度,并进行对比。

[0119] 分类精度=测试集中正确分类的样本/测试集样本总数

[0120] 表1.两种方法分类精度一览表

[0121]	仿真实验方法	随机森林	本方法
	分类精度	75.43%	92.93%

[0122] 表1中分别列出了本发明和现有随机森林技术的分类精度计算结果,从表1可见,本发明的分类精度为92.93%,这个指标远高于随机森林分类方法,证明本发明可以得到比随机森林方法更高的步态能量图的分类精度。

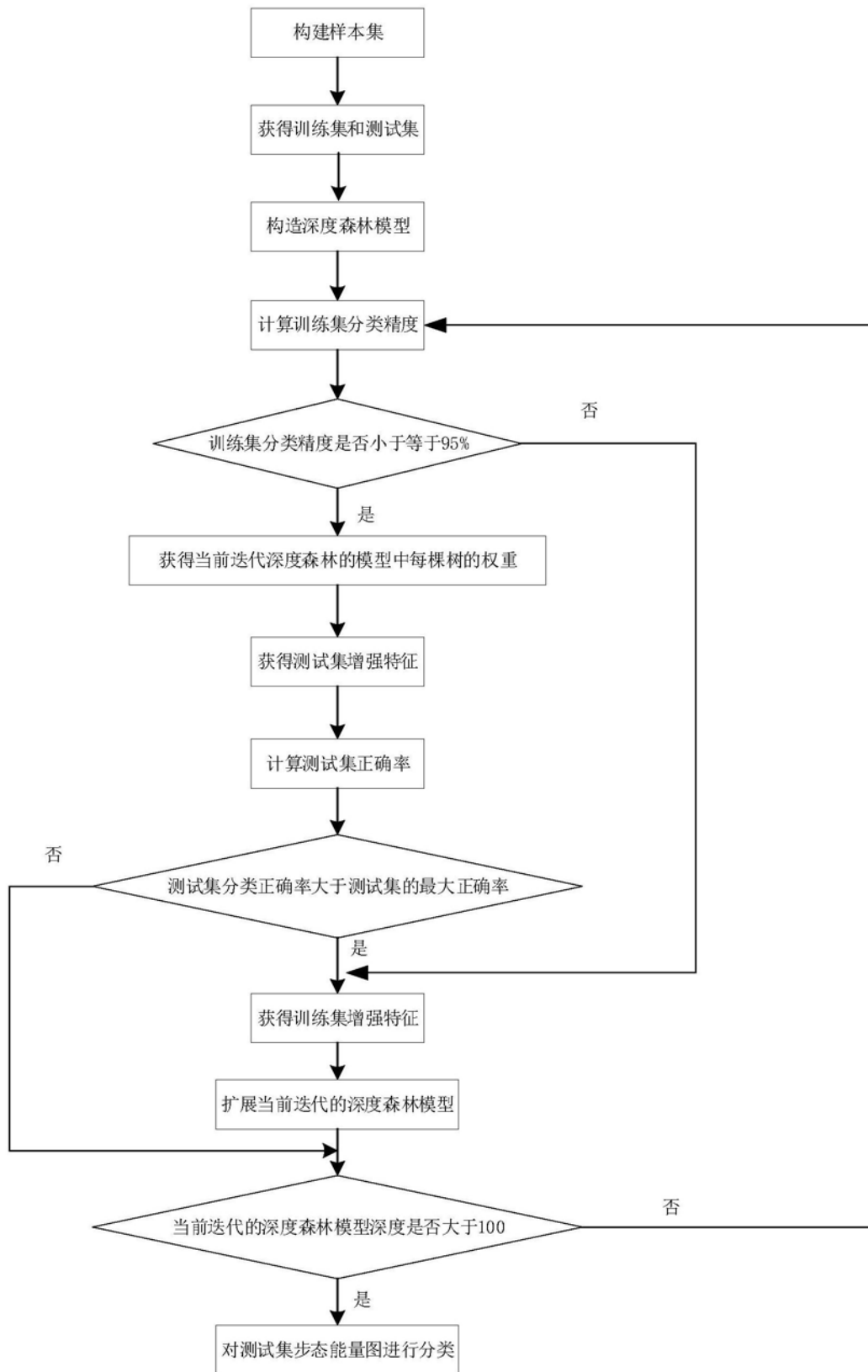


图1