



(12)发明专利申请

(10)申请公布号 CN 111414895 A

(43)申请公布日 2020.07.14

(21)申请号 202010280653.5

G06N 3/08(2006.01)

(22)申请日 2020.04.10

(71)申请人 上海卓繁信息技术股份有限公司
地址 200030 上海市徐汇区番禺路1028号
202室

(72)发明人 刘凤余 张琼 王培基

(74)专利代理机构 上海容慧专利代理事务所
(普通合伙) 31287

代理人 于晓菁

(51) Int. Cl.

G06K 9/00(2006.01)

G06K 9/34(2006.01)

G06K 9/46(2006.01)

G06K 9/62(2006.01)

G06N 3/04(2006.01)

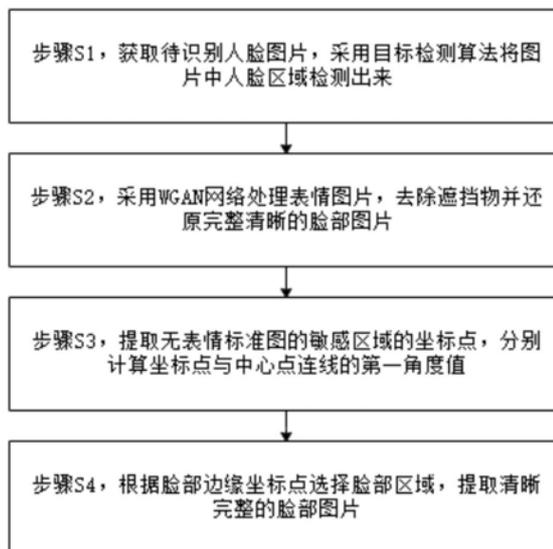
权利要求书1页 说明书5页 附图1页

(54)发明名称

一种人脸识别方法,装置及存储设备

(57)摘要

本发明提供了一种人脸识别方法,装置及存储设备,可在复杂环境下准确获取脸部图像,包括获取待识别人脸图片,采用目标检测算法将图片中人脸区域检测出来;采用WGAN网络处理表情图片,去除遮挡物并还原完整清晰的脸部图片;采用FCN网络处理脸部图片,将脸部与头部的其它部分区分,并记录脸部边缘坐标点;根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸部图片。



1. 一种人脸识别方法,其特征在于,包括:
获取待识别人脸图片,采用目标检测算法将图片中人脸区域检测出来;
采用WGAN网络处理表情图片,去除遮挡物并还原完整清晰的脸部图片;
采用FCN网络处理脸部图片,将脸部与头部的其它部分区分,并记录脸部边缘坐标点;
根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸部图片。
2. 如权利要求1所述的一种人脸识别方法,其特征在于,使用基于SSD网络的人脸检测算法获取图像中的人脸位置,将图像输入SSD网络,从网络的输出中获得人脸区域的坐标。
3. 如权利要求1所述的一种人脸识别方法,其特征在于,
所述WGAN网络模型包括第一生成网络和第一判别网络,所述第一生成网络包括卷积层,空洞卷积层,归一化层,平均池化层,lrelu激活层,relu激活层和残差块,所述第一判别网络包括5个卷积块。
4. 如权利要求3所述的一种人脸识别方法,其特征在于,
采用WGAN网络模型对有遮挡物和无遮挡物的脸部图片数据集进行学习,还原完整清晰的脸部图片,所述遮挡物包括眼镜、胡须、帽子、额头遮挡物等。
5. 如权利要求1所述的一种人脸识别方法,其特征在于,
所述FCN网络采用反卷积层对卷积结果进行采样,对脸部图片的像素作出预测,从而进行像素级别的分类以完成图像的分割。
6. 如权利要求1所述的一种人脸识别方法,其特征在于,
根据所述脸部边缘坐标点选择人脸区域,去除人脸区域以外的背景区域并对人脸区域做标准化处理。
7. 一种人脸识别装置,其特征在于,包括:
人脸图像获取模块,用以获取待识别的人脸图片;
遮挡物去除模块,用以通过WGAN网络根据现有的表情图片生成无表情的标准图;
脸部区域区分模块,用以采用FCN网络处理脸部图片,将脸部与头部的其它部分区分,并记录脸部边缘坐标点;
脸部图片提取模块,用以根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸部图片。
8. 一种存储装置,其特征在于,所述存储介质存储有多条指令,所述指令适于处理器进行加载,以执行权利要求1至6任一项所述的满意度测评方法中的步骤。

一种人脸识别方法,装置及存储设备

技术领域

[0001] 本发明涉及机器学习技术领域,具体涉及一种人脸识别方法,装置及存储设备。

背景技术

[0002] 随着“放、管、服”改革的持续推进,为了推进政务窗口服务质量的提升,需要借助人工智能技术以对工作人员的服务做一个评价参考。

[0003] 因此,采用基于深度学习的图像识别技术对顾客进行表情识别,以根据顾客表情对服务进行评分是一种可行的思路。但是制约表情识别效果提升的因素之一,就是复杂场景下顾客脸部检测不够准确,如各种遮挡物,背景等极易影响识别效果。

发明内容

[0004] 本发明的目的是提供一种人脸识别方法,装置及存储设备,用以在复杂环境下准确获取脸部图像。

[0005] 为了达到上述目的,本发明一方面提供一种人脸识别方法,包括:

[0006] 获取待识别人脸图片,采用目标检测算法将图片中人脸区域检测出来;

[0007] 采用WGAN网络处理表情图片,去除遮挡物并还原完整清晰的脸部图片;

[0008] 采用FCN网络处理脸部图片,将脸部与头部的其它部分区分,并记录脸部边缘坐标点;

[0009] 根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸部图片。

[0010] 进一步的,使用基于SSD网络的人脸检测算法获取图像中的人脸位置,将图像输入SSD网络,从网络的输出中获得人脸区域的坐标。

[0011] 进一步的,所述WGAN网络模型包括第一生成网络和第一判别网络,所述第一生成网络包括卷积层,空洞卷积层,批归一化层,平均池化层,lrelu激活层,relu激活层和残差块,所述第一判别网络包括5个卷积块。

[0012] 进一步的,采用WGAN网络模型对有遮挡物和无遮挡物的脸部图片数据集进行学习,还原完整清晰的脸部图片,所述遮挡物包括眼镜、胡须、帽子、额头遮挡物等。

[0013] 进一步的,所述FCN网络采用反卷积层对卷积结果进行采样,从而对脸部图片的像素作出预测,从而进行像素级别的分类以完成图像的分割。

[0014] 进一步的,根据所述脸部边缘坐标点选择人脸区域,去除人脸区域以外的背景区域并对人脸区域做标准化处理。

[0015] 另一方面,本发明还提供一种人脸识别装置,包括:

[0016] 人脸图像获取模块,用以获取待识别的人脸图片;

[0017] 遮挡物去除模块,用以通过WGAN网络根据现有的表情图片生成无表情的标准图;

[0018] 脸部区域区分模块,用以采用FCN网络处理脸部图片,将脸部与头部的其它部分区分,并记录脸部边缘坐标点;

[0019] 脸部图片提取模块,用以根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸

部图片。

[0020] 另一方面,本发明还提供一种存储装置,所述存储介质存储有多条指令,所述指令适于处理器进行加载,以执行权利要求上述人脸识别方法中的步骤。

[0021] 本发明提供了一种人脸识别方法,装置及存储设备,可在复杂环境下准确获取脸部图像,包括获取待识别人脸图片,采用目标检测算法将图片中人脸区域检测出来;采用WGAN网络处理表情图片,去除遮挡物并还原完整清晰的脸部图片;采用FCN网络处理脸部图片,将脸部与头部的其它区域区分,并记录脸部边缘坐标点;根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸部图片。

附图说明

[0022] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其它的附图。

[0023] 图1是本发明一个实施例的一种人脸识别方法的方法流程图。

[0024] 图2为本发明一个实施例的一种人脸识别装置的系统架构图。

具体实施方式

[0025] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0026] 在以下的描述中,涉及到“一些实施例”,其描述了所有可能实施例的子集,但是可以理解,“一些实施例”可以是所有可能实施例的相同子集或不同子集,并且可以在不冲突的情况下相互结合。

[0027] 如果申请文件中出现“第一/第二”的类似描述则增加以下的说明,在以下的描述中,所涉及的术语“第一\第二\第三”仅仅是是区别类似的对象,不代表针对对象的特定排序,可以理解地,“第一\第二\第三”在允许的情况下可以互换特定的顺序或先后次序,以使这里描述的本申请实施例能够以除了在这里图示或描述的以外的顺序实施。

[0028] 除非另有定义,本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同。本文中所使用的术语只是为了描述本申请实施例的目的,不是旨在限制本申请。

[0029] 下面参照附图描述根据本发明实施例提出的一种人脸识别方法,装置和存储设备,首先将参照附图描述根据本发明实施例提出的一种人脸识别方法。

[0030] 图1是本发明一个实施例的人脸识别方法的方法流程图。如图1所示,该测评方法包括以下步骤:

[0031] 步骤S1,获取待识别人脸图片,采用目标检测算法将图片中人脸区域检测出来。

[0032] 作为一个实施例,本发明使用基于SSD网络的人脸检测算法获取图像中的人脸位置,将图像输入SSD网络,从网络的输出中获得人脸区域的坐标。

[0033] 具体的,SSD算法(Single Shot MultiBox Detector),SSD的骨干网络是基于传统的图像分类网络,例如VGG,ResNet等。在本实施例经过卷积层(con.layer)和池化层(max pooling)的处理,得到人脸的特征图(feature maps)。在这些特征图上进行回归计算,得到人脸的位置和类别,从而获得人脸区域的坐标。

[0034] 步骤S2,采用WGAN网络处理表情图片,去除遮挡物并还原完整清晰的脸部图片;

[0035] 生成式对抗网络(Generative Adversarial Network,GAN)是一种生成式深度学习网络模型,近年来在计算机视觉领域应用效果良好。随着深度学习及移动设备的快速发展,将深度学习应用在图像处理、图像生成、图像风格迁移等领域有非常巨大的应用效果。GAN能生成目标数据集,以弥补训练数据不足的缺陷,因此对深度学习意义重大。

[0036] 在一个实施例中,所述WGAN网络模型包括第一生成网络和第一判别网络,所述第一生成网络包括卷积层,空洞卷积层,批归一化层,平均池化层,lrelu激活层,relu激活层和残差块,所述第一判别网络包括5个卷积块。

[0037] WGAN网络的原理为:

[0038] (1) 训练第一代第一生成网络和第一代第一判别网络。从一个噪声分布从采样作为输入,输入第一代第一生成器网络,它能生成一些很差的脸部图片。然后第一代第一判别网络可以将现有的脸部图片与生成的脸部图片进行判断,判别器可以准确的判别出哪些是现有的脸部图片哪些是生成的脸部图片。

[0039] (2) 训练第二代生成网络和第二代判别网络。可以生成更好的脸部图片,让第二代第一判别器很难判断哪些是更好的脸部图片,哪些是更差的脸部图片。

[0040] (3) 不断有第3代,第4代,。。。第n代。直到第n代判别网络几乎无法判断出更好的脸部图片时停止。此时的第n代生成器就是最好的脸部图片的生成器,可用于脸部图片生成。

[0041] 具体的,本实施例的WGAN网络模型的训练过程为:加载数据,数据归一化,数据重塑;构造沃瑟斯坦距离,作为判别网络的评判指标;从高斯分布中采样出噪声图片X作为生成网络G的输入,又经过全连接构成的网络训练,将生成器网络的输出作为判别网络D的一个输入,真实训练样本作为另一个输入,经过判别网络进行损失函数的计算,最后通过最小化损失函数的大小,得到深度卷积神经网络识别要求的图像。

[0042] 本实施例的的WGAN网络模型的损失函数为:

$$[0043] \quad W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} E_{(x,y) \sim \gamma} [\|x - y\|]$$

[0044] 其中, $\Pi(P_r, P_g)$ 是Pr和Pg组合起来的所有可能的联合分布的集合,即 $\Pi(P_r, P_g)$ 中每一个分布的边缘分布都是Pr和Pg,对于每一个可能的联合分布 γ 而言,可以从中采样 $(x, y) \sim \gamma$ 得到一个真实样本x和一个生成样本y,并算出这对样本之间的距离 $\|x - y\|$,得到该联合分布 γ 下样本对距离的期望值 $E_{(x,y) \sim \gamma} [\|x - y\|]$,从所有可能的联合分布中获取这个期望值所取到的下界。

[0045] 具体的,本实施例分别采用2000张有遮挡物和无遮挡物的图像对WGAN网络模型进行训练,学习到将有遮挡物图转变为无遮挡物的图像特征。

[0046] 步骤S3,提取无表情标准图的敏感区域的坐标点,分别计算坐标点与中心点连线的第一角度值。

[0047] 在一个实施例中,无表情标准图的敏感区域坐标点为15个,包括每只眼睛的上下

眼皮中心点,每条眉毛两端的端点,上下嘴唇的中心点,嘴角两端的端点,每只眼睛两端眼角的端点。

[0048] 进一步的,计算鼻子中心点与敏感区域坐标点之间连线的角度值,分别生成14个无表情标准图的第一角度值(X_1, X_2, \dots, X_{14})。

[0049] 步骤S3,采用FCN网络处理脸部图片,将脸部与头部的其它部分区分,并记录脸部边缘坐标点。

[0050] 本领域技术人员应该可以理解,FCN网络对图像进行像素级的分类,从而解决了语义级别的图像分割(semantic segmentation)问题。与经典的CNN在卷积层之后使用全连接层得到固定长度的特征向量进行分类(全联接层+softmax输出)不同,FCN网络可以接受任意尺寸的输入图像,采用反卷积层对最后一个卷积层的feature map进行上采样,使它恢复到输入图像相同的尺寸,从而可以对每个像素都产生了一个预测,同时保留了原始输入图像中的空间信息,最后在上采样的特征图上进行逐像素分类。

[0051] 步骤S4,根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸部图片。

[0052] 具体的,根据所述脸部边缘坐标点选择人脸区域,去除人脸区域以外的背景区域并对人脸区域做标准化处理,生成标准化的脸部图片。

[0053] 如图2所示,另一方面,本发明还提供一种人脸识别装置,包括:

[0054] 人脸图像获取模块101,用以获取待识别的人脸图片;

[0055] 遮挡物去除模块102,用以通过WGAN网络根据现有的表情图片生成无表情的标准图;

[0056] 脸部区域区分模块103,用以采用FCN网络处理脸部图片,将脸部与头部的其它部分区分,并记录脸部边缘坐标点;

[0057] 脸部图片提取模块104,用以根据脸部边缘坐标点选择脸部区域,提取清晰完整的脸部图片。

[0058] 另一方面,本发明还提供一种存储装置,所述存储介质存储有多条指令,所述指令适于处理器进行加载,以执行权利要求上述人脸识别方法中的步骤。

[0059] 另一方面,本发明还提供一种存储装置所述存储介质存储有多条指令,所述指令适于处理器进行加载,以执行权利要求上述满意度测评方法中的步骤。

[0060] 本发明实施例的计算机存储介质,可以采用一个或多个计算机可读的介质的任意组合。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPR0M或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本文件中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0061] 计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可

读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0062] 计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括——但不限于无线、电线、光缆、RF等等,或者上述的任意合适的组合。

[0063] 可以以一种或多种程序设计语言或其组合来编写用于执行本发明操作的计算机程序代码,所述程序设计语言包括面向对象的程序设计语言—诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或终端上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0064] 尽管上面已经示出和描述了本发明的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本发明的限制,本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

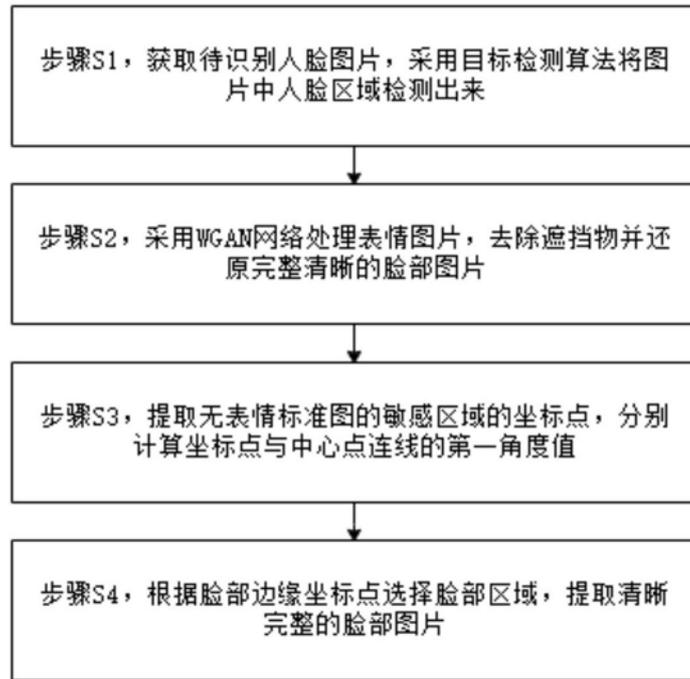


图1



图2