



(12) 发明专利

(10) 授权公告号 CN 109948441 B

(45) 授权公告日 2021.03.26

(21) 申请号 201910115532.2

(22) 申请日 2019.02.14

(65) 同一申请的已公布的文献号  
申请公布号 CN 109948441 A

(43) 申请公布日 2019.06.28

(73) 专利权人 北京奇艺世纪科技有限公司  
地址 100080 北京市海淀区北一街2号鸿城  
拓展大厦10、11层

(72) 发明人 刘思阳

(74) 专利代理机构 北京润泽恒知识产权代理有  
限公司 11319

代理人 莎日娜

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

(56) 对比文件

WO 0237401 A3, 2002.07.11

CN 106203533 A, 2016.12.07

CN 108875540 A, 2018.11.23

CN 107679490 A, 2018.02.09

审查员 甘宇

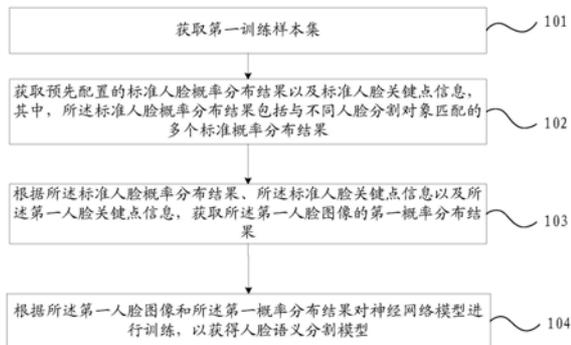
权利要求书4页 说明书15页 附图5页

(54) 发明名称

模型训练、图像处理方法和、装置、电子设备及  
计算机可读存储介质

(57) 摘要

本发明提供了一种模型训练、图像处理方  
法、装置、电子设备、计算机可读存储介质,该模  
型训练方法包括:获取第一训练样本集,第一训  
练样本集包括第一人脸图像以及第一人脸图像  
的第一人脸关键点信息;获取预先配置的标准人  
脸概率分布结果以及标准人脸关键点信息,其  
中,标准人脸概率分布结果包括与不同人脸分割  
对象匹配的多个标准概率分布结果;根据标准人  
脸概率分布结果、标准人脸关键点信息以及第  
一人脸关键点信息,获取第一人脸图像的第一  
概率分布结果;根据第一人脸图像和第一概率  
分布结果对神经网络模型进行训练,以获得人  
脸语义分割模型。本发明训练得到的人脸语义  
分割模型提升了对人脸图像的不同人脸语义  
区域的分割准确度。



1. 一种模型训练方法,其特征在于,包括:

获取第一训练样本集,所述第一训练样本集包括第一人脸图像以及所述第一人脸图像的第一人脸关键点信息;

获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人脸关键点信息,获取所述第一人脸图像的第一概率分布结果,所述第一概率分布结果表征在所述第一人脸图像中每个分割对象对应的区域内,各个坐标点对应的像素属于所述人脸分割对象的概率;

根据所述第一人脸图像和所述第一概率分布结果对神经网络模型进行训练,以获得人脸语义分割模型。

2. 根据权利要求1所述的方法,其特征在于,所述获取第一训练样本集之前,所述方法还包括:

获取第二训练样本集,所述第二训练样本集包括第二人脸图像、与不同分割对象匹配的多个第二人脸语义分割结果、第二人脸关键点信息,其中,所述多个第二人脸语义分割结果以及所述第二人脸关键点信息均与所述第二人脸图像匹配;

针对所述第二训练样本集中的任意一个第二人脸图像,根据所述第二人脸关键点信息和预先配置的标准人脸关键点信息,获取第一仿射矩阵;

根据所述第一仿射矩阵,对所述多个第二人脸语义分割结果分别进行仿射变换,生成与所述不同人脸分割对象匹配的多个标准人脸语义分割结果;

根据每个第二人脸图像对应的所述多个标准人脸语义分割结果,获取与不同人脸分割对象匹配的多个标准概率分布结果。

3. 根据权利要求2所述的方法,其特征在于,所述根据每个第二人脸图像对应的所述多个标准人脸语义分割结果,获取与不同人脸分割对象匹配的多个标准概率分布结果,包括:

按照如下公式获取与人脸分割对象1匹配的标准概率分布结果img\_{a,b}^{(l)};

$$img_{a,b}^{(l)} = \frac{1}{k} \sum_k img_{a,b}^{(l)j};$$

其中,j表示所述第二训练样本集中的第j个第二人脸图像, $1 \leq j \leq k$ ,j为整数,j分别取值1,2,3……k,k为所述第二训练样本集中所述第二人脸图像的总个数;

$img^{(1)j}$ 表示与人脸分割对象1匹配的标准人脸语义分割结果,其中,所述标准语义分割结果对应于第j个第二人脸图像;

$img_{a,b}^{(1)j}$ 表示所述标准人脸语义分割结果img^{(1)j}中坐标(a,b)的数值,其中,坐标(a,b)为所述标准人脸语义分割结果img^{(1)j}中的每个坐标;

$img_{a,b}^{(l)}$ 为k个第二人脸图像对应的k个img\_{a,b}^{(l)j}的平均值, $img_{a,b}^{(l)}$ 表示坐标(a,b)对应的像素点属于所述人脸分割对象1的概率值。

4. 根据权利要求1所述的方法,其特征在于,所述根据所述标准人脸概率分布结果、所

述标准人脸关键点信息以及所述第一人脸关键点信息,获取所述第一人脸图像的第一概率分布结果,包括:

根据所述标准人脸关键点信息和所述第一人脸关键点信息,获取第二仿射矩阵;

根据所述第二仿射矩阵,对所述多个标准概率分布结果分别进行仿射变换,得到所述第一人脸图像的第一概率分布结果,其中,所述第一概率分布结果包括与所述不同人脸分割对象匹配的多个第一局部概率分布结果。

5. 根据权利要求1所述的方法,其特征在于,所述第一概率分布结果包括第一概率分布矩阵,所述根据所述第一人脸图像和所述第一概率分布结果对神经网络模型进行训练,包括:

获取与所述第一人脸图像匹配的第一人脸图像矩阵;

对所述第一人脸图像矩阵和所述第一概率分布矩阵进行矩阵连接处理;

根据经过所述矩阵连接处理后的矩阵数据对神经网络模型进行训练。

6. 一种图像处理方法,其特征在于,包括:

获取待分割的第一人脸图像;

获取所述第一人脸图像的第一人脸关键点信息;

获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人脸关键点信息,获取所述第一人脸图像的第一概率分布结果,所述第一概率分布结果表征在所述第一人脸图像中每个分割对象对应的区域内,各个坐标点对应的像素属于所述人脸分割对象的概率;

将所述第一人脸图像和所述第一概率分布结果输入至预先经过训练的人脸语义分割模型,得到与不同人脸分割对象匹配的多个目标人脸语义分割结果。

7. 一种模型训练装置,其特征在于,包括:

第一获取模块,用于获取第一训练样本集,所述第一训练样本集包括第一人脸图像以及所述第一人脸图像的第一人脸关键点信息;

第二获取模块,用于获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

第三获取模块,用于根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人脸关键点信息,获取所述第一人脸图像的第一概率分布结果,所述第一概率分布结果表征在所述第一人脸图像中每个分割对象对应的区域内,各个坐标点对应的像素属于所述人脸分割对象的概率;

训练模块,用于根据所述第一人脸图像和所述第一概率分布结果对神经网络模型进行训练,以获得人脸语义分割模型。

8. 根据权利要求7所述的装置,其特征在于,所述装置还包括:

第四获取模块,用于获取第二训练样本集,所述第二训练样本集包括第二人脸图像、与不同分割对象匹配的多个第二人脸语义分割结果、第二人脸关键点信息,其中,所述多个第二人脸语义分割结果以及所述第二人脸关键点信息均与所述第二人脸图像匹配;

第五获取模块,用于针对所述第二训练样本集中的任意一个第二人脸图像,根据所述第二人脸关键点信息和预先配置的标准人脸关键点信息,获取第一仿射矩阵;

生成模块,用于根据所述第一仿射矩阵,对所述多个第二人脸语义分割结果分别进行仿射变换,生成与所述不同人脸分割对象匹配的多个标准人脸语义分割结果;

第六获取模块,用于根据每个第二人脸图像对应的所述多个标准人脸语义分割结果,获取与不同人脸分割对象匹配的多个标准概率分布结果。

9. 根据权利要求8所述的装置,其特征在于,

所述第六获取模块,还用于按照如下公式获取与人脸分割对象1匹配的标准概率分布结果

$$img_{a,b}^{(l)}; img_{a,b}^{(l)} = \frac{1}{k} \sum_k img_{a,b}^{(l)j};$$

其中,j表示所述第二训练样本集中的第j个第二人脸图像, $1 \leq j \leq k$ ,j为整数,j分别取值1,2,3……k,k为所述第二训练样本集中所述第二人脸图像的总个数;

$img^{(1)j}$ 表示与人脸分割对象1匹配的标准人脸语义分割结果,其中,所述标准语义分割结果对应于第j个第二人脸图像;

$img_{a,b}^{(1)j}$ 表示所述标准人脸语义分割结果 $img^{(1)j}$ 中坐标(a,b)的数值,其中,坐标(a,b)为所述标准人脸语义分割结果 $img^{(1)j}$ 中的每个坐标;

$img_{a,b}^{(l)}$ 为k个第二人脸图像对应的k个 $img_{a,b}^{(1)j}$ 的平均值, $img_{a,b}^{(l)}$ 表示坐标(a,b)对应的像素点属于所述人脸分割对象1的概率值。

10. 根据权利要求7所述的装置,其特征在于,所述第三获取模块包括:

第一获取子模块,用于根据所述标准人脸关键点信息和所述第一人脸关键点信息,获取第二仿射矩阵;

变换子模块,用于根据所述第二仿射矩阵,对所述多个标准概率分布结果分别进行仿射变换,得到所述第一人脸图像的第一概率分布结果,其中,所述第一概率分布结果包括与所述不同人脸分割对象匹配的多个第一局部概率分布结果。

11. 根据权利要求7所述的装置,其特征在于,所述训练模块包括:

第二获取子模块,用于当所述第一概率分布结果包括第一概率分布矩阵时,获取与所述第一人脸图像匹配的第一人脸图像矩阵;

处理子模块,用于对所述第一人脸图像矩阵和所述第一概率分布矩阵进行矩阵连接处理;

训练子模块,用于根据经过所述矩阵连接处理后的矩阵数据对神经网络模型进行训练。

12. 一种图像处理装置,其特征在于,包括:

第一获取模块,用于获取待分割的第一人脸图像;

第二获取模块,用于获取所述第一人脸图像的第一人脸关键点信息;

第三获取模块,用于获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

第四获取模块,用于根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人臉关键点信息,获取所述第一人臉图像的第一概率分布结果,所述第一概率分布结果表征在所述第一人臉图像中每个分割对象对应的区域内,各个坐标点对应的像素属于所述人脸分割对象的概率;

语义分割模块,用于将所述第一人臉图像和所述第一概率分布结果输入至预先经过训练的人脸语义分割模型,得到与不同人脸分割对象匹配的多个目标人脸语义分割结果。

13. 一种电子设备,其特征在于,包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的模型训练程序或图像处理程序,所述模型训练程序被所述处理器执行时实现如权利要求1至5中任一项所述的模型训练方法的步骤,所述图像处理程序被所述处理器执行时实现如权利要求6所述的图像处理方法的步骤。

14. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有模型训练程序或图像处理程序,所述模型训练程序被处理器执行时实现如权利要求1至5中任一项所述的模型训练方法中的步骤,所述图像处理程序被所述处理器执行时实现如权利要求6所述的图像处理方法的步骤。

## 模型训练、图像处理方法、装置、电子设备及计算机可读存储介质

### 技术领域

[0001] 本发明涉及机器学习技术领域,特别是涉及一种模型训练、图像处理方法、装置、电子设备及计算机可读存储介质。

### 背景技术

[0002] 随着计算机技术的不断发展,对人脸图像进行美化的软件越来越多。在对人脸图像进行美化时,需要识别人脸图像中对应不同人脸部件(例如五官、脸部轮廓)的区域,其中,不同人脸部件的区域可以理解为具有不同人脸语义的区域,为了识别人脸图像中具有不同人脸语义的区域,需要进行人脸图像进行语义分割。

[0003] 目前的人脸语义分割算法主要是将人脸图像作为训练数据,通过对神经网络模型的训练,来使得经过训练的神经网络模型可以对人脸图像进行语义分割,从而得到对应不同人脸部件的区域。

[0004] 但是,相关技术中训练得到的神经网络模型容易对输入的人脸图像进行错误的语义分割,即将不属于某个人脸部件的多个像素点分割到该人脸部件对应的区域中,存在着人脸部件的分割结果不准确的问题。

### 发明内容

[0005] 本发明提供了一种模型训练、图像处理方法、装置、电子设备及计算机可读存储介质,以解决相关技术中的人脸语义分割方案在对人脸图像进行语义分割时,所存在的人脸部件的分割结果准确度较低的问题。

[0006] 为了解决上述问题,根据本发明的第一方面,本发明公开了一种模型训练方法,包括:

[0007] 获取第一训练样本集,所述第一训练样本集包括第一人脸图像以及所述第一人脸图像的第一人脸关键点信息;

[0008] 获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

[0009] 根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人脸关键点信息,获取所述第一人脸图像的第一概率分布结果;

[0010] 根据所述第一人脸图像和所述第一概率分布结果对神经网络模型进行训练,以获得人脸语义分割模型。

[0011] 根据本发明的第二方面,本发明公开了一种图像处理方法,包括:

[0012] 获取待分割的第一人脸图像;

[0013] 获取所述第一人脸图像的第一人脸关键点信息;

[0014] 获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

[0015] 根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人臉关键点信息,获取所述第一人臉图像的第一概率分布结果;

[0016] 将所述第一人臉图像和所述第一概率分布结果输入至预先经过训练的人臉语义分割模型,得到与不同人臉分割对象匹配的多个目标人臉语义分割结果。

[0017] 根据本发明的第三方面,本发明还公开了一种模型训练装置,包括:

[0018] 第一获取模块,用于获取第一训练样本集,所述第一训练样本集包括第一人臉图像以及所述第一人臉图像的第一人臉关键点信息;

[0019] 第二获取模块,用于获取预先配置的标准人臉概率分布结果以及标准人臉关键点信息,其中,所述标准人臉概率分布结果包括与不同人臉分割对象匹配的多个标准概率分布结果;

[0020] 第三获取模块,用于根据所述标准人臉概率分布结果、所述标准人臉关键点信息以及所述第一人臉关键点信息,获取所述第一人臉图像的第一概率分布结果;

[0021] 训练模块,用于根据所述第一人臉图像和所述第一概率分布结果对神经网络模型进行训练,以获得人臉语义分割模型。

[0022] 根据本发明的第四方面,本发明还公开了一种图像处理装置,包括:

[0023] 第一获取模块,用于获取待分割的第一人臉图像;

[0024] 第二获取模块,用于获取所述第一人臉图像的第一人臉关键点信息;

[0025] 第三获取模块,用于获取预先配置的标准人臉概率分布结果以及标准人臉关键点信息,其中,所述标准人臉概率分布结果包括与不同人臉分割对象匹配的多个标准概率分布结果;

[0026] 第四获取模块,用于根据所述标准人臉概率分布结果、所述标准人臉关键点信息以及所述第一人臉关键点信息,获取所述第一人臉图像的第一概率分布结果;

[0027] 语义分割模块,用于将所述第一人臉图像和所述第一概率分布结果输入至预先经过训练的人臉语义分割模型,得到与不同人臉分割对象匹配的多个目标人臉语义分割结果。

[0028] 根据本发明的第五方面,本发明还公开了一种电子设备,包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的模型训练程序或图像处理程序,所述模型训练程序被所述处理器执行时实现如上述任意一项所述的模型训练方法的步骤,所述图像处理程序被所述处理器执行时实现上述图像处理方法的步骤。

[0029] 根据本发明的第六方面,本发明还公开了一种计算机可读存储介质,所述计算机可读存储介质上存储有模型训练程序或图像处理程序,所述模型训练程序被处理器执行时实现如上述任意一项所述的模型训练方法中的步骤,所述图像处理程序被所述处理器执行时实现上述图像处理方法的步骤。

[0030] 与现有技术相比,本发明包括以下优点:

[0031] 在本发明实施例中,利用先验的标准人臉关键点信息、预先配置的标准人臉概率分布结果以及第一训练集中每个第一人臉图像的人臉关键点信息,可以获取到第一人臉图像的第一概率分布结果,由于第一概率分布结果是基于标准人臉概率分布结果和标准人臉关键点信息而确定出来的,其中,标准人臉概率分布结果在标准人臉图像中能够对各个人臉分割对象所在的区别进行限制,因此,该第一概率分布结果也能够对第一人臉图像中各

个人脸分割对象所在的区域进行限制,那么在利用第一人脸图像以及第一人脸图像的第一概率分布结果来对神经网络模型进行训练时,则可以使训练得到的人脸语义分割模型学习在该第一概率分布结果限制的各个区域内进行各个人脸分割对象的语义识别和分割,能够防止训练后得到的人脸语义分割模型对各个区域外的部分进行误识别,提升了训练后得到的人脸语义分割模型对人脸图像的不同人脸语义区域的分割准确度。

### 附图说明

- [0032] 图1是本发明的一种模型训练方法实施例的步骤流程图;
- [0033] 图2是本发明的一种标准人脸图像实施例的示意图;
- [0034] 图3是本发明的另一种模型训练方法实施例的步骤流程图;
- [0035] 图4是本发明的一种第二人脸图像实施例的示意图;
- [0036] 图5是本发明的一种获取与鼻子匹配的标准概率分布结果实施例的示意图;
- [0037] 图6是本发明的一种图像处理方法实施例的步骤流程图;
- [0038] 图7是一种模型训练装置实施例的结构框图;
- [0039] 图8是一种图像处理装置实施例的结构框图。

### 具体实施方式

[0040] 为使本发明的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0041] 参照图1,示出了本发明的一种模型训练方法实施例的步骤流程图,具体可以包括如下步骤:

[0042] 步骤101,获取第一训练样本集;

[0043] 本发明实施例为了训练神经网络模型(其网络结构可以是任意一种神经网络模型,本发明对此不做限制)来得到人脸语义分割模型,这里需要获取该第一训练样本集,其中,所述第一训练样本集可以包括多组样本,每一组样本包括人脸图像(即第一人脸图像),以及该人脸图像的人脸关键点信息(即第一人脸关键点信息);

[0044] 其中,第一人脸关键点信息可以是多个预设人脸关键点的坐标信息,其中,预设人脸关键点可以包括但不限于左眼关键点、右眼关键点、鼻子关键点、嘴巴关键点。

[0045] 以左眼为例,左眼关键点的坐标信息可以是第一人脸图像中的左眼上的任意一个坐标位置,本发明实施例中,左眼关键点的坐标信息可以是左眼的瞳孔的坐标信息。对于其他预设人脸关键点的坐标信息也是类似的,这里不再列举。

[0046] 例如人脸图像样本*i*的5个预设人脸关键点的坐标信息分别为:

[0047] 左眼关键点的坐标为:  $p_1^i(x_1^i, y_1^i)$ ; 右眼关键点的坐标为:  $p_2^i(x_2^i, y_2^i)$ ; 鼻尖关键点的坐标为:  $p_3^i(x_3^i, y_3^i)$ ; 左嘴角关键点的坐标为:  $p_4^i(x_4^i, y_4^i)$ ; 右嘴角关键点的坐标为:  $p_4^i(x_4^i, y_4^i)$ 。

[0048] 步骤102,获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

[0049] 其中,本发明实施例的方法可以预先配置标准人脸关键点信息,标准人脸关键点

信息表达了标准人脸图像(图像中的人脸完整、人脸中的各个人脸部件都是正对镜头的,例如如图2所示的标准人脸图像)中多个预设人脸关键点的坐标信息。其中,图2示出了5个标准人脸关键点,分别为左眼关键点21、右眼关键点22、鼻尖关键点23、左嘴角关键点24和右嘴角关键点25。

[0050] 对于上述5个标准人脸关键点的坐标信息,本发明实施例的方法可以根据先验知识来人工配置,本步骤只需要获取预先配置的这5个标准人脸关键点的坐标信息。

[0051] 例如,左眼关键点21的坐标为: $p_1(x_1, y_1)$ ;右眼关键点22的坐标为: $p_2(x_2, y_2)$ ;鼻尖关键点23的坐标为: $p_3(x_3, y_3)$ ;左嘴角关键点24的坐标为: $p_4(x_4, y_4)$ ;右嘴角关键点25的坐标为: $p_5(x_5, y_5)$ 。

[0052] 此外,本发明实施例的方法还可以预先配置标准人脸概率分布结果,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果。

[0053] 其中,如图2所示,人脸分割对象可以包括人脸图像中的人脸部件以及人脸部件的背景31,在本发明实施例中,人脸分割对象可以包括但不限于头发32、面部33、左眉毛34、右眉毛35、左眼睛36、右眼睛37、鼻子38、上嘴唇39、口腔39、下嘴唇39(这里为了简便,嘴部的三个分割对象都以一个图像区域示出)。因此,一般情况下,一个人脸图像可以包括11个人脸分割对象,本发明实施例训练神经网络模型的目的,旨在使得训练后得到人脸语义分割模型能够对人脸图像分割出11个人脸分割对象。

[0054] 其中,本发明实施例的方法可以预先配置标准人脸概率分布结果,该标准人脸概率分布结果为基于人脸语义分割的数据集(即下述第二训练数据集)而确定的结果。其中,标准人脸概率分布结果表达了在基于分割数据集所确定的标准人脸图像(例如图2所示的图像)中,每个人脸分割对象对应的区域内的各个坐标点对应的像素属于该人脸分割对象的概率。因此,所述标准人脸概率分布结果可以包括与不同人脸分割对象匹配的多个标准概率分布结果。

[0055] 本步骤需要获取预先配置的标准人脸概率分布结果。

[0056] 步骤103,根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人眼关键点信息,获取所述第一人眼图像的第一概率分布结果;

[0057] 可选地,在执行步骤103时,可以首先根据所述标准人脸关键点信息和所述第一人眼关键点信息,获取第二仿射矩阵;然后,根据所述第二仿射矩阵,对所述多个标准概率分布结果分别进行仿射变换,得到所述第一人眼图像的第一概率分布结果。

[0058] 具体而言,对于第一训练集中的任意一组样本,可以根据所述标准人脸关键点信息和所述第一人眼关键点信息,获取第二仿射矩阵。

[0059] 以第一训练集中的人脸图像1为例,可以对5个标准人脸关键点 $p_1 \sim p_5$ 的坐标和图像样本1的5个预设人脸关键点 $p_1^1 \sim p_5^1$ 的坐标进行仿射变换,得到两组人脸关键点之间的仿射矩阵 $A^1$ :

$$[0060] \quad A^1 = \begin{bmatrix} a_1 & a_2 & t_x \\ a_3 & a_4 & t_y \\ 0 & 0 & 1 \end{bmatrix};$$

[0061] 其中,仿射矩阵中的a、t表示向量。

[0062] 而多个标准概率分布结果表达了标准人脸图像中每个分割对象对应的区域内,各

个坐标点对应的像素属于该人脸分割对象的概率,而第一训练集中的第一人脸图像并非标准人脸图像,其脸部可以是任意角度,那么为了获取在第一人脸图像中每个分割对象对应的区域内,各个坐标点对应的像素属于该人脸分割对象的概率,这里可以根据仿射矩阵 $A^1$ ,对所述多个标准概率分布结果分别进行仿射变换,从而可以得到所述第一人脸图像的第一概率分布结果。该第一概率分布结果则表达了在第一人脸图像中每个分割对象对应的区域内,各个坐标点对应的像素属于该人脸分割对象的概率。那么概率高的像素点属于对应人脸分割对象的可能性就大,这样,第一概率分布结果相当于表达了第一人脸图像中任意一个像素点属于某个人脸分割对象的概率,且一个像素点只属于一个人脸分割对象。

[0063] 那么如果按照不同的人脸分割对象来说,则所述第一概率分布结果可以包括与所述不同人脸分割对象匹配的多个第一局部概率分布结果。

[0064] 例如,对应于鼻子的局部概率分布结果包括:对应于鼻子的区域1中各个像素点属于鼻子的概率。

[0065] 这样,本发明实施例根据训练样本集中第一人脸图像的第一人脸关键点信息和标准人脸关键点信息,获取它们之间的仿射矩阵,再利用仿射矩阵来对标准人脸的多个标准概率分布结果分别进行仿射变换,从而可以得到该第一人脸图像的第一概率分布结果,其中,所述第一概率分布结果可以包括与所述不同人脸分割对象匹配的多个第一局部概率分布结果,这样,就获得了训练样本集中第一人脸图像中各个人脸分割对象的大致边界,那么再利用该第一概率分布结果来对神经网络模型进行训练时,则可以使得训练后得到的人脸语义分割模型学习到在各个人脸分割对象的大致边界内进行各个人脸分割对象的识别,从而得到与不同人脸分割对象匹配的多个个人脸语义分割结果,提升人脸语义分割结果的准确度。

[0066] 步骤104,根据所述第一人脸图像和所述第一概率分布结果对神经网络模型进行训练,以获得人脸语义分割模型。

[0067] 在本发明实施例中,对于第一训练集中的每个第一人脸图像都执行步骤102~步骤104的训练步骤,从而可以使得经过训练后的神经网络模型,即得到的所述人脸语义分割模型能够用于根据待分割人脸图像的概率分布结果(与上述第一概率分布结果表达的意思类似),对待分割人脸图像进行人脸语义的细化分割,输出与不同人脸分割对象匹配的多个目标人脸语义分割结果。

[0068] 在本发明实施例中,利用先验的标准人脸关键点信息、预先配置的标准人脸概率分布结果以及第一训练集中每个第一人脸图像的人脸关键点信息,可以获取到第一人脸图像的第一概率分布结果,由于第一概率分布结果是基于标准人脸概率分布结果和标准人脸关键点信息而确定出来的,其中,标准人脸概率分布结果在标准人脸图像中能够对各个人脸分割对象所在的区别进行限制,因此,该第一概率分布结果也能够对第一人脸图像中各个人脸分割对象所在的区域进行限制,那么在利用第一人脸图像以及第一人脸图像的第一概率分布结果来对神经网络模型进行训练时,则可以使训练得到的人脸语义分割模型学习在该第一概率分布结果限制的各个区域内进行各个人脸分割对象的语义识别和分割,能够防止训练后得到的人脸语义分割模型对各个区域外的部分进行误识别,提升了训练后得到的人脸语义分割模型对人脸图像的不同人脸语义区域的分割准确度。

[0069] 相较于相关技术中只将图像的RGB(红绿蓝)矩阵送入神经网络模型进行训练方案

来说,本发明实施例的方法将第一人脸图像以及该第一人脸图像的概率分布结果(即该第一人脸图像的多个个人脸分割对象的概率分布矩阵)均输入至神经网络模型进行训练,可以使得训练得到的人脸语义分割模型更加精准的对人脸分割对象进行准确的语义分割。

[0070] 可选地,所述第一概率分布结果可以包括第一概率分布矩阵,那么在执行步骤104时,可以首先获取与所述第一人脸图像匹配的第一人脸图像矩阵;然后,对所述第一人脸图像矩阵和所述第一概率分布矩阵进行矩阵连接处理;最后,根据经过所述矩阵连接处理后的矩阵数据对神经网络模型进行训练。

[0071] 具体而言,在对第一训练集中的图像样本经过步骤101~步骤103的处理后,可以对每个第一人脸图像都得到对应的第一概率分布结果,而第一概率分布结果可以以矩阵的形式存在,这里记为第一概率分布矩阵。此外,由于第一概率分布结果包括与不同人脸分割对象匹配的多个第一局部概率分布结果,例如人脸分割对象的数量为上述列举的11个人脸分割对象,那么这里的第一概率分布矩阵则是 $W*H*11$ 的概率分布矩阵。每一个 $W*H$ 的概率分布矩阵对应一个人脸分割对象。

[0072] 其中,第一人脸图像为RGB图像,因此,第一人脸图像中的每个像素点都包括R(红色)值、G(绿色)值和B(蓝色)值,例如第一人脸图像的尺寸为 $W*H$ ,即宽度为 $W$ ,在宽度方向上包括 $W$ 个像素点,长度为 $H$ ,在长度方向上包括 $H$ 个像素点。那么第一人脸图像中任意一个颜色都可以构成一个 $W*H*1$ 的矩阵,第一人脸图像中的R值构成一个矩阵,G值构成一个矩阵,B值构成一个矩阵,且每个矩阵的长度是 $H$ ,宽度是 $W$ ,所以,第一人脸图像的第一人脸图像矩阵是 $W*H*3$ 的图像矩阵,即包括三层 $W*H$ 的矩阵。

[0073] 因此,一个图像的图像矩阵就是以矩阵的方式来表达出的图像数据,或者说是该图像的矩阵结构。

[0074] 那么可以将所述第一人脸图像矩阵 $W*H*3$ 和所述第一概率分布矩阵 $W*H*11$ 作矩阵连接处理,得到 $W*H*14$ 的矩阵数据;最后,根据 $W*H*14$ 的矩阵数据对神经网络模型进行训练。

[0075] 这样,本发明实施例通过获取第一训练集中第一人脸图像的第一人脸图像矩阵,以矩阵数据的方式来表达图像数据,此外,还获取到以矩阵数据的方式来表达的该第一人脸图像中,与不同人脸分割对象匹配的第一概率分布矩阵,并将这两个矩阵数据作拼接,利用拼接后的矩阵数据来对神经网络模型进行训练,相对于相关技术中只将训练样本的图像矩阵输入到神经网络模型进行人脸语义的分割,本发明实施例还将该训练样本的与不同人脸分割对象匹配的第一概率分布矩阵输入到神经网络模型进行训练,可以使得训练后得到的人脸语义分割模型更加精准的对各个人脸部件进行语义分割,而不会将不属于某个人脸分割对象的像素点分割到该人脸分割对象对应的区域内,提升了训练后的人脸语义分割模型对人脸语义分割的准确度。

[0076] 可选地,在步骤101之前,根据本发明实施例的方法还可以包括:生成标准人脸关键点信息。

[0077] 在生成标准人脸关键点信息时,本发明实施例的方法可以根据经验人工设置在例如如图2所示的标准人脸图像的状态下,各个预设人脸关键点的标准坐标信息,例如在图2中,左眼关键点21的坐标为: $p_1(x_1, y_1)$ ;右眼关键点22的坐标为: $p_2(x_2, y_2)$ ;鼻尖关键点23的坐标为: $p_3(x_3, y_3)$ ;左嘴角关键点24的坐标为: $p_4(x_4, y_4)$ ;右嘴角关键点25的坐标为: $p_5(x_5,$

$y_5)$ 。

[0078] 可选地,在步骤101之前,根据本发明实施例的方法还可以包括:生成标准人脸概率分布结果。

[0079] 在生成标准人脸概率分布结果时,本发明实施例还可以通过人脸语义分割的数据集,来统计标准人脸图像中各个人脸分割对象的概率分布,即标准人脸概率分布结果。

[0080] 这样,本发明实施例通过生成标准人脸关键点的信息,以及标准人脸概率分布结果,并利用这两项信息,来对第一训练集中的第一人脸图像进行处理,来获取第一人脸图像的第一概率分布结果,并利用第一人脸图像和第一概率分布结果来对神经网络模型进行训练,使得训练得到的人脸语义分割模型能够利用第一概率分布结果来对待分割人脸图像中的目标人脸分割对象所在的区域进行约束和限制,防止将目标人脸分割对象所在的区域外面的像素点分割到该目标人脸分割对象的情况,提升了人脸语义分割的准确度。

[0081] 可选地,在生成标准人脸概率分布结果时,可以通过如图3所示的方法来实现:

[0082] S201,获取第二训练样本集;

[0083] 其中,所述第二训练样本集包括多组样本;

[0084] 每一组样本包括一个第二人脸图像、与不同分割对象匹配的多个第二人脸语义分割结果、第二人脸关键点信息,其中,所述多个第二人脸语义分割结果以及所述第二人脸关键点信息均与所述第二人脸图像匹配;

[0085] 在一个示例中,以一组样本为例来说明,该一组样本包括一个原始人脸图像,即上述第二人脸图像,该原始人脸图像对应11个人工分割图片(这11个分割图片分别对应于上文所述的11个人脸分割对象),即上述多个第二人脸语义分割结果,一个文本文件。

[0086] 其中,每张分割图片表示对应人脸分割对象的人工分割结果,且每张分割图片是二值化处理后的图片,在该分割图片中属于该分割对象的像素点取值为1,不属于该分割对象的像素值取值为0。可以理解为每张分割图片都是一个人脸分割对象的矩阵。

[0087] 其中,文本文件中记录有该原始人脸图像中5个预设人脸关键点的坐标信息,例如左眼,右眼,鼻尖,左嘴角,右嘴角五个关键点的坐标。

[0088] 需要说明的是,本发明实施例中各个人脸图像的预设关键点的坐标信息和标准人脸关键点的坐标信息是基于同一个直角坐标系的,且二者的关键点类型相同。

[0089] 在一个示例中,图4示出了一个第二人脸图像的11个分割对象对应的区域,以及第二人脸图像中的5个预设人脸关键点,其中,图4中的附图标记与图2中的附图标记相同,因此,这里不再赘述。

[0090] S202,针对所述第二训练样本集中的任意一个第二人脸图像,根据所述第二人脸关键点信息和预先配置的标准人脸关键点信息,获取第一仿射矩阵;

[0091] 举例来说,图4所示的第二人脸图像中的5个预设人脸关键点具有5个坐标信息,而根据经验确定的5个标准人脸关键点也具有5个坐标信息,那么可以对这两组人脸关键点的坐标信息进行仿射变换,从而得到第一仿射矩阵。

[0092] S203,根据所述第一仿射矩阵,对所述多个第二人脸语义分割结果分别进行仿射变换,生成与所述不同人脸分割对象匹配的多个标准人脸语义分割结果;

[0093] 继续以上例举例来说,第二训练集中的第二人脸图像对应有11个人脸语义分割结果,那么按照上述第一仿射矩阵,来对11个人脸语义分割结果分别进行仿射变换,即图4中

的11个人脸分割对象对应的11个区域,向图2所示的11个人脸分割对象对应的11个区域进行仿射变换,从而可以将第二训练样本集中第二人脸图像的角度不标准的11个区域变换为角度标准的11个区域,即上述多个标准人脸语义分割结果。

[0094] S204,根据每个第二人脸图像对应的所述多个标准人脸语义分割结果,获取与不同人脸分割对象匹配的多个标准概率分布结果。

[0095] 其中,第二训练集中的每个第二人脸图像都可以经过S202~S203来得到角度标准的11个区域,那么可以利用第二训练集中所有第二人脸图像的多组11个区域,来获取在标准人脸图像中的标准人脸概率分布结果,其中,该标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果。

[0096] 其中,多个标准概率分布结果表达了标准人脸图像中每个分割对象对应的区域内,各个坐标点对应的像素属于该人脸分割对象的概率。

[0097] 这样,本发明实施例为了生成标准人脸概率分布结果,利用标准人脸关键点信息和第二训练样本集中第二人脸图像的第二人脸关键点信息,来获取第一仿射矩阵,并利用该第一仿射矩阵,来对第二人脸图像的人工分割的多个人脸语义分割结果进行仿射变换,从而能够将第二人脸图像中原本角度不正的多个人脸语义分割结果转换成在标准人脸状态下,即角度标准下的多个标准人脸语义分割结果,最后,利用第二训练集中每个第二人脸图像对应的所述多个标准人脸语义分割结果,来获取与不同人脸分割对象匹配的多个标准概率分布结果,多个标准概率分布结果表达了标准人脸图像中每个分割对象对应的区域内,各个坐标点对应的像素属于该人脸分割对象的概率。这样,就通过人脸语义分割的数据集(即第二训练集)统计出了与不同人脸分割对象匹配的多个标准概率分布结果,其中,第二训练集中的第二人脸图像中越可能属于某个人脸分割对象的像素点,其概率值就越大。那么后续在利用该标准人脸概率分布结果来对神经网络模型进行训练时,可以对各个分割目标(即人脸分割对象)进行边界的约束,防止在人脸分割对象所在区域的外面出现错误分割的情况。

[0098] 可选地,在执行S204时,可以按照如下公式1获取与人脸分割对象1匹配的标准概率分布结果img\_{a,b}^{(l)};

[0099] 
$$img_{a,b}^{(l)} = \frac{1}{k} \sum_k img_{a,b}^{(l)j}, \text{ 公式 1};$$

[0100] 其中,j表示所述第二训练样本集中的第j个第二人脸图像, $1 \leq j \leq k$ ,j为整数,j分别取值1,2,3……k,k为所述第二训练样本集中所述第二人脸图像的总个数;

[0101]  $img^{(1)j}$ 表示与人脸分割对象1匹配的标准人脸语义分割结果,其中,所述标准语义分割结果对应于第j个第二人脸图像;

[0102]  $img_{a,b}^{(l)j}$ 表示所述标准人脸语义分割结果img^{(1)j}中坐标(a,b)的数值,其中,坐标(a,b)为所述标准人脸语义分割结果img^{(1)j}中的每个坐标;

[0103]  $img_{a,b}^{(l)}$ 为k个第二人脸图像对应的k个img\_{a,b}^{(l)j}的平均值, $img_{a,b}^{(l)}$ 表示坐标(a,b)对应的像素点属于所述人脸分割对象1的概率值。

[0104] 这里结合图5以人脸分割对象1为鼻子,来对本发明实施例的上述公式1进行解释:

[0105] 例如 $k=3$ ,即第二训练样本集包括3组样本,图5示出了第二人脸图像1中鼻子的标准人脸语义分割结果51,第二人脸图像2中鼻子的标准人脸语义分割结果52和第二人脸图像3中鼻子的标准人脸语义分割结果53,其中,标准人脸语义分割结果也是一个矩阵,矩阵中的数值为1表示该位置对应的像素点属于鼻子,矩阵中的数值为0表示该位置对应的像素点不属于鼻子。

[0106] 那么为了利用图5所示的标准人脸语义分割结果51、标准人脸语义分割结果52和标准人脸语义分割结果53,来获取与鼻子匹配的标准概率分布结果54,需要对矩阵中的四个坐标的值重新进行计算。如图5所示,在计算标准概率分布结果54的左上角的坐标(1,1)的数值时,则是将第二人脸图像1中鼻子的标准人脸语义分割结果51中的左上角的坐标(1,1)的数值、第二人脸图像2中鼻子的标准人脸语义分割结果52中的左上角的坐标(1,1)的数值、以及第二人脸图像3中鼻子的标准人脸语义分割结果53中的左上角的坐标(1,1)的数值作平均值计算,即 $(0+1+1)/3=0.67$ ,得到标准概率分布结果54中左上角的坐标(1,1)对应的像素点属于鼻子的概率为0.67。对于标准概率分布结果54中其他坐标上的概率值的计算采用类似的方式,具体参照图5即可。这样,就根据第二训练集得到了人脸分割对象——鼻子的标准概率分布结果,即以概率矩阵的方式来表达出在标准人脸图像中像素点属于鼻子的概率。类似的,还可以得到其他人脸分割对象的标准概率分布结果。

[0107] 这样,本发明实施例根据第二训练集中每个第二人脸图像对应的多个标准人脸语义分割结果,按照人脸分割对象的不同,来分别计算各个人脸分割对象下的标准概率分布结果,从而得到与不同人脸分割对象匹配的多个标准概率分布结果,通过配置该多个标准概率分布结果,可以使用该多个标准概率分布结果来对神经网络模型进行训练,从而达到对各个人脸分割对象所在的区域进行限制的目的,使得训练后得到的人脸语义分割模型只在各个人脸分割对象所在的大区域内进行细化分割,而不会将不属于该人脸分割对象的像素点分割到该人脸分割对象下,提升了人脸语义分割的准确度。

[0108] 这里示意性的示出了本发明实施例的神经网络模型,也即训练后得到的人脸语义分割模型的网络结构:

[0109] 网络输入为 $(w*h*14)$ 的矩阵;

[0110] 第1层为有128个 $3\times 3$ 卷积核的卷积层,输入尺寸为 $w\times h\times 14$ ,输出尺寸为 $w\times h\times 128$ 。

[0111] 第2层为有128个 $3\times 3$ 卷积核的卷积层,输入尺寸为 $w\times h\times 128$ ,输出尺寸为 $w\times h\times 128$ 。

[0112] 第3层为有 $2\times 2$ 池化核的最大池化层,输入尺寸为 $w\times h\times 128$ ,输出尺寸为 $w/2\times h/2\times 128$ 。

[0113] 第4层为有256个 $3\times 3$ 卷积核的卷积层,输入尺寸为 $w/2\times h/2\times 128$ ,输出尺寸为 $w/2\times h/2\times 256$ 。

[0114] 第5层为有256个 $3\times 3$ 卷积核的卷积层,输入尺寸为 $w/2\times h/2\times 256$ ,输出尺寸为 $w/2\times h/2\times 256$ 。

[0115] 第6层为有 $2\times 2$ 池化核的最大池化层,输入尺寸为 $w/2\times h/2\times 256$ ,输出尺寸为 $w/4\times h/4\times 256$ 。

[0116] 第7层为有512个 $3\times 3$ 卷积核的卷积层,输入尺寸为 $w/4\times h/4\times 256$ ,输出尺寸为 $w/$

$4 \times h/4 \times 512$ 。

[0117] 第8层为有512个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/4 \times h/4 \times 512$ ,输出尺寸为 $w/4 \times h/4 \times 512$ 。

[0118] 第9层为有 $2 \times 2$ 池化核的最大池化层,输入尺寸为 $w/4 \times h/4 \times 512$ ,输出尺寸为 $w/8 \times h/8 \times 512$ 。

[0119] 第10层为有1024个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/8 \times h/8 \times 512$ ,输出尺寸为 $w/8 \times h/8 \times 1024$ 。

[0120] 第11层为有1024个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/8 \times h/8 \times 1024$ ,输出尺寸为 $w/8 \times h/8 \times 1024$ 。

[0121] 第12层为有 $2 \times 2$ 池化核的最大池化层,输入尺寸为 $w/8 \times h/8 \times 1024$ ,输出尺寸为 $w/16 \times h/16 \times 1024$ 。

[0122] 第13层为有2048个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/16 \times h/16 \times 1024$ ,输出尺寸为 $w/16 \times h/16 \times 2048$ 。

[0123] 第14层为有2048个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/16 \times h/16 \times 2048$ ,输出尺寸为 $w/16 \times h/16 \times 2048$ 。

[0124] 第15层为有 $2 \times 2$ 池化核的最大池化层,输入尺寸为 $w/16 \times h/16 \times 2048$ ,输出尺寸为 $w/32 \times h/32 \times 2048$ 。

[0125] 第16层为行和列上采样因子为(2,2)的上采样层,输入尺寸为 $w/32 \times h/32 \times 2048$ ,输出尺寸为 $w/16 \times h/16 \times 2048$ 。

[0126] 第17层为有1024个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/16 \times h/16 \times 2048$ ,输出尺寸为 $w/16 \times h/16 \times 1024$ 。

[0127] 第18层为拼接层,将第17层的输出和第12层的输出拼接,输入尺寸为两个 $w/16 \times h/16 \times 1024$ ,输出尺寸为 $w/16 \times h/16 \times 2048$ 。

[0128] 第19层为有1024个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/16 \times h/16 \times 2048$ ,输出尺寸为 $w/16 \times h/16 \times 1024$ 。

[0129] 第20层为行和列上采样因子为(2,2)的上采样层,输入尺寸为 $w/16 \times h/16 \times 1024$ ,输出尺寸为 $w/8 \times h/8 \times 1024$ 。

[0130] 第21层为有512个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/8 \times h/8 \times 1024$ ,输出尺寸为 $w/8 \times h/8 \times 512$ 。

[0131] 第22层为拼接层,将第21层的输出和第9层的输出拼接,输入尺寸为两个 $w/8 \times h/8 \times 512$ ,输出尺寸为 $w/8 \times h/8 \times 1024$ 。

[0132] 第23层为有512个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/8 \times h/8 \times 1024$ ,输出尺寸为 $w/8 \times h/8 \times 512$ 。

[0133] 第24层为行和列上采样因子为(2,2)的上采样层,输入尺寸为 $w/8 \times h/8 \times 512$ ,输出尺寸为 $w/4 \times h/4 \times 512$ 。

[0134] 第25层为有256个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/4 \times h/4 \times 512$ ,输出尺寸为 $w/4 \times h/4 \times 256$ 。

[0135] 第26层为拼接层,将第25层的输出和第6层的输出拼接,输入尺寸为两个 $w/4 \times h/4 \times 256$ ,输出尺寸为 $w/4 \times h/4 \times 512$ 。

[0136] 第27层为有256个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/4 \times h/4 \times 512$ ,输出尺寸为 $w/4 \times h/4 \times 256$ 。

[0137] 第28层为行和列上采样因子为(2,2)的上采样层,输入尺寸为 $w/4 \times h/4 \times 256$ ,输出尺寸为 $w/2 \times h/2 \times 256$ 。

[0138] 第29层为有128个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/2 \times h/2 \times 256$ ,输出尺寸为 $w/2 \times h/2 \times 128$ 。

[0139] 第30层为拼接层,将第29层的输出和第3层的输出拼接,输入尺寸为两个 $w/2 \times h/2 \times 128$ ,输出尺寸为 $w/2 \times h/2 \times 256$ 。

[0140] 第31层为有128个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w/2 \times h/2 \times 256$ ,输出尺寸为 $w/2 \times h/2 \times 128$ 。

[0141] 第32层为行和列上采样因子为(2,2)的上采样层,输入尺寸为 $w/2 \times h/2 \times 128$ ,输出尺寸为 $w \times h \times 128$ 。

[0142] 第33层为有128个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w \times h \times 128$ ,输出尺寸为 $w \times h \times 128$ 。

[0143] 第34层为拼接层,将第33层的输出和第1层的输出拼接,输入尺寸为两个 $w \times h \times 128$ ,输出尺寸为 $w \times h \times 256$ 。

[0144] 第35层为有128个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w \times h \times 256$ ,输出尺寸为 $w \times h \times 128$ 。

[0145] 第36层为有311个 $3 \times 3$ 卷积核的卷积层,输入尺寸为 $w \times h \times 128$ ,输出尺寸为 $w \times h \times 114$ 。

[0146] 参照图6,示出了本发明的一种图像处理方法实施例的步骤流程图,具体可以包括如下步骤:

[0147] 步骤601,获取待分割的第三人脸图像;

[0148] 步骤602,获取所述第三人脸图像的第三人脸关键点信息;

[0149] 这里的第三人脸关键点信息与上述模型训练方法中的第一人脸关键点信息的原理类似,这里不再赘述。

[0150] 步骤603,获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

[0151] 其中,本步骤与上述模型训练方法中的步骤102类似,具体参照上文,这里不再赘述。

[0152] 步骤604,根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第三人脸关键点信息,获取所述第三人脸图像的第三概率分布结果;

[0153] 其中,本步骤的执行原理与上述模型训练方法中的步骤103类似,具体参照上文,这里不再赘述。

[0154] 步骤605,将所述第三人脸图像和所述第三概率分布结果输入至预先经过训练的人脸语义分割模型,得到与不同人脸分割对象匹配的多个目标人脸语义分割结果;

[0155] 其中,这里的人脸语义分割模型即为上述实施例训练得到的人脸语义分割模型。

[0156] 其中,所述人脸语义分割模型用于根据所述第一概率分布结果,对第一人脸图像进行人脸语义的细化分割,输出与不同人脸分割对象匹配的多个目标人脸语义分割结果。

[0157] 在本发明实施例中,利用先验的标准人脸关键点信息、预先配置的标准人脸概率分布结果以及待分割的第三人脸图像的第三人脸关键点信息,可以获取到第三人脸图像的第三概率分布结果,该第三概率分布结果能够对第三人脸图像中各个人脸分割对象所在的区域进行限制,那么在将第三人脸图像以及第三人脸图像的第三概率分布结果输入至预先经过训练的人脸语义分割模型后,人脸语义分割模型则可以在该第三概率分布结果限制的各个区域内对第三人脸图像进行各个人脸分割对象的语义识别和分割,能够防止人脸语义分割模型对各个区域外的部分进行误识别,提升了对人脸图像进行人脸语义分割时的准确度。

[0158] 需要说明的是,对于方法实施例,为了简单描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本发明实施例并不受所描述的动作顺序的限制,因为依据本发明实施例,某些步骤可以采用其他顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作并不一定是本发明实施例所必须的。

[0159] 与上述本发明实施例所提供的模型训练方法相对应,参照图7,示出了本发明一种模型训练装置实施例的结构框图,具体可以包括如下模块:

[0160] 第一获取模块701,用于获取第一训练样本集,所述第一训练样本集包括第一人脸图像以及所述第一人脸图像的第三人脸关键点信息;

[0161] 第二获取模块702,用于获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息,其中,所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

[0162] 第三获取模块703,用于根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人脸关键点信息,获取所述第一人脸图像的第一概率分布结果;

[0163] 训练模块704,用于根据所述第一人脸图像和所述第一概率分布结果对神经网络模型进行训练,以获得人脸语义分割模型。

[0164] 可选地,所述装置还包括:

[0165] 第四获取模块,用于获取第二训练样本集,所述第二训练样本集包括第二人脸图像、与不同分割对象匹配的多个第二人脸语义分割结果、第二人脸关键点信息,其中,所述多个第二人脸语义分割结果以及所述第二人脸关键点信息均与所述第二人脸图像匹配;

[0166] 第五获取模块,用于针对所述第二训练样本集中的任意一个第二人脸图像,根据所述第二人脸关键点信息和预先配置的标准人脸关键点信息,获取第一仿射矩阵;

[0167] 生成模块,用于根据所述第一仿射矩阵,对所述多个第二人脸语义分割结果分别进行仿射变换,生成与所述不同人脸分割对象匹配的多个标准人脸语义分割结果;

[0168] 第六获取模块,用于根据每个第二人脸图像对应的所述多个标准人脸语义分割结果,获取与不同人脸分割对象匹配的多个标准概率分布结果。

[0169] 可选地,所述第六获取模块,还用于按照如下公式获取与人脸分割对象1匹配的标准概率分布结果 $img_{a,b}^{(l)}$ ;

$$[0170] \quad img_{a,b}^{(l)} = \frac{1}{k} \sum_k img_{a,b}^{(l)j};$$

[0171] 其中,  $j$ 表示所述第二训练样本集中的第 $j$ 个第二人脸图像,  $1 \leq j \leq k$ ,  $j$ 为整数,  $j$ 分别取值 $1, 2, 3, \dots, k$ ,  $k$ 为所述第二训练样本集中所述第二人脸图像的总个数;

[0172]  $img^{(1)j}$ 表示与人脸分割对象 $l$ 匹配的标准人脸语义分割结果, 其中, 所述标准语义分割结果对应于第 $j$ 个第二人脸图像;

[0173]  $img_{a,b}^{(l)j}$ 表示所述标准人脸语义分割结果 $img^{(1)j}$ 中坐标 $(a, b)$ 的数值, 其中, 坐标 $(a, b)$ 为所述标准人脸语义分割结果 $img^{(1)j}$ 中的每个坐标;

[0174]  $img_{a,b}^{(l)}$ 为 $k$ 个第二人脸图像对应的 $k$ 个 $img_{a,b}^{(l)j}$ 的平均值,  $img_{a,b}^{(l)}$ 表示坐标 $(a, b)$ 对应的像素点属于所述人脸分割对象 $l$ 的概率值。

[0175] 可选地, 所述第三获取模块703包括:

[0176] 第一获取子模块, 用于根据所述标准人脸关键点信息和所述第一人脸关键点信息, 获取第二仿射矩阵;

[0177] 变换子模块, 用于根据所述第二仿射矩阵, 对所述多个标准概率分布结果分别进行仿射变换, 得到所述第一人脸图像的第一概率分布结果, 其中, 所述第一概率分布结果包括与所述不同人脸分割对象匹配的多个第一局部概率分布结果。

[0178] 可选地, 所述训练模块704包括:

[0179] 第二获取子模块, 用于当所述第一概率分布结果包括第一概率分布矩阵时, 获取与所述第一人脸图像匹配的第一人脸图像矩阵;

[0180] 处理子模块, 用于对所述第一人脸图像矩阵和所述第一概率分布矩阵进行矩阵连接处理;

[0181] 训练子模块, 用于根据经过所述矩阵连接处理后的矩阵数据对神经网络模型进行训练。

[0182] 对于模型训练装置实施例而言, 由于其与模型训练方法实施例基本相似, 所以描述的比较简单, 相关之处参见方法实施例的部分说明即可。

[0183] 与上述本发明实施例所提供的图像处理方法相对应, 参照图8, 示出了本发明一种图像处理装置实施例的结构框图, 具体可以包括如下模块:

[0184] 第七获取模块801, 用于获取待分割的第一人脸图像;

[0185] 第八获取模块802, 用于获取所述第一人脸图像的第一人脸关键点信息;

[0186] 第九获取模块803, 用于获取预先配置的标准人脸概率分布结果以及标准人脸关键点信息, 其中, 所述标准人脸概率分布结果包括与不同人脸分割对象匹配的多个标准概率分布结果;

[0187] 第十获取模块804, 用于根据所述标准人脸概率分布结果、所述标准人脸关键点信息以及所述第一人脸关键点信息, 获取所述第一人脸图像的第一概率分布结果;

[0188] 语义分割模块805, 用于将所述第一人脸图像和所述第一概率分布结果输入至预先经过训练的人脸语义分割模型, 得到与不同人脸分割对象匹配的多个目标人脸语义分割结果。

[0189] 对于图像处理装置实施例而言, 由于其与图像处理方法实施例基本相似, 所以描述的比较简单, 相关之处参见方法实施例的部分说明即可。

[0190] 根据本发明的又一个实施例, 本发明还提供了一种电子设备, 包括: 存储器、处理

器及存储在所述存储器上并可在所述处理器上运行的模型训练程序或图像处理程序,所述模型训练程序被所述处理器执行时实现如上述任意一个实施例所述的模型训练方法的步骤,所述图像处理程序被所述处理器执行时实现上述任意一个实施例所述的图像处理方法的步骤。

[0191] 根据本发明的再一个实施例,本发明还提供了一种计算机可读存储介质,所述计算机可读存储介质上存储有模型训练程序或图像处理程序,所述模型训练程序被处理器执行时实现如上述任意一个实施例所述的模型训练方法中的步骤,所述图像处理程序被所述处理器执行时实现上述任意一个实施例所述的图像处理方法的步骤。

[0192] 本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。

[0193] 本领域内的技术人员应明白,本发明实施例的实施例可提供为方法、装置、或计算机程序产品。因此,本发明实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0194] 本发明实施例是参照根据本发明实施例的方法、终端设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理终端设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理终端设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0195] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理终端设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0196] 这些计算机程序指令也可装载到计算机或其他可编程数据处理终端设备上,使得在计算机或其他可编程终端设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程终端设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0197] 尽管已描述了本发明实施例的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本发明实施例范围的所有变更和修改。

[0198] 最后,还需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者终端设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者终端设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要

素,并不排除在包括所述要素的过程、方法、物品或者终端设备中还存在另外的相同要素。

[0199] 以上对本发明所提供的一种模型训练方法、一种模型训练装置、一种图像处理方法、一种图像处理装置、一种电子设备、一种计算机可读存储介质,进行了详细介绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

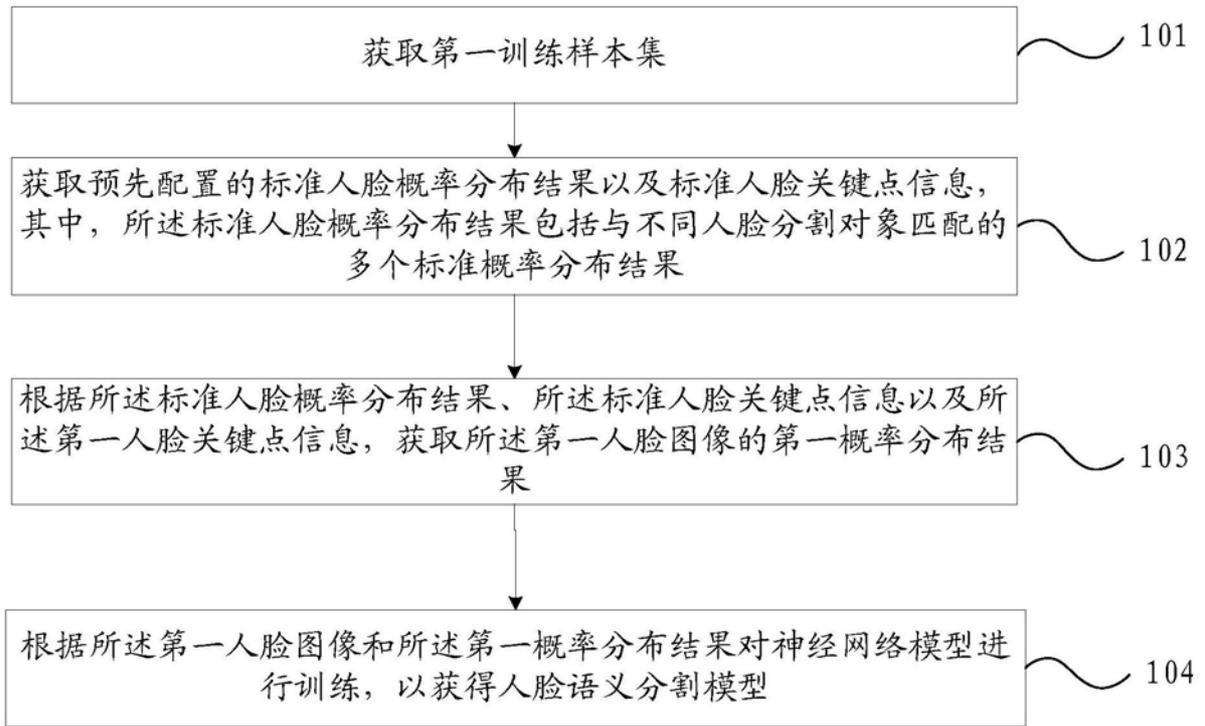


图1

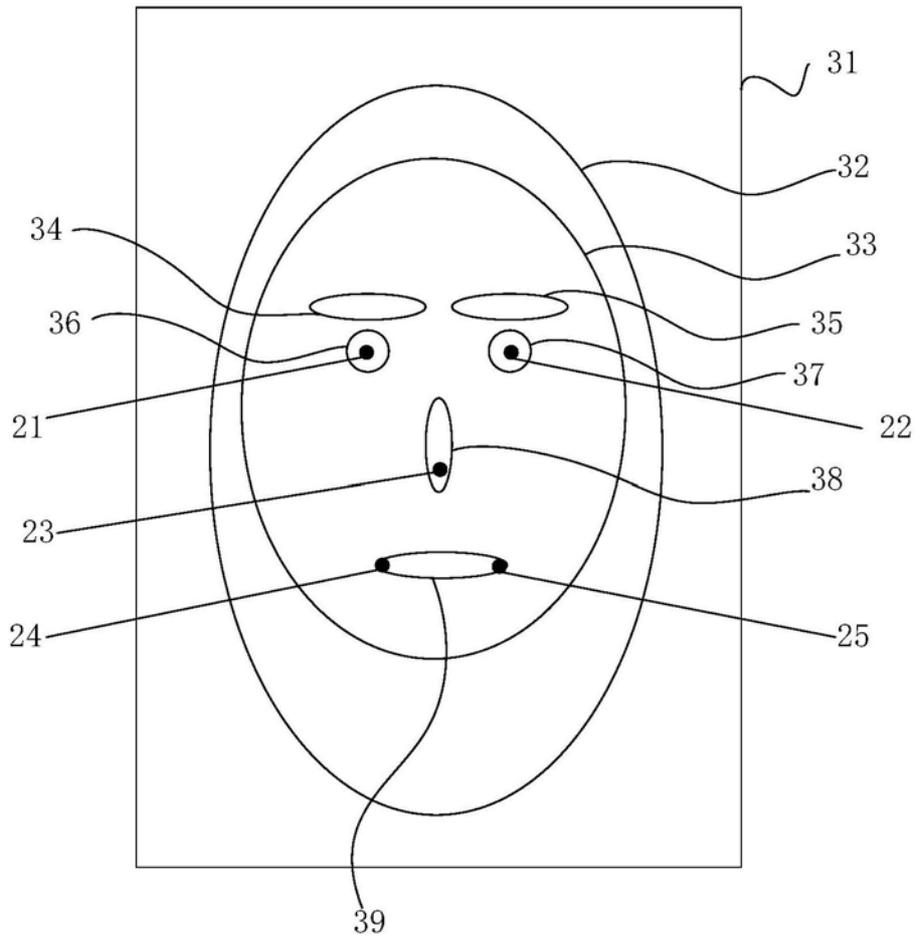


图2

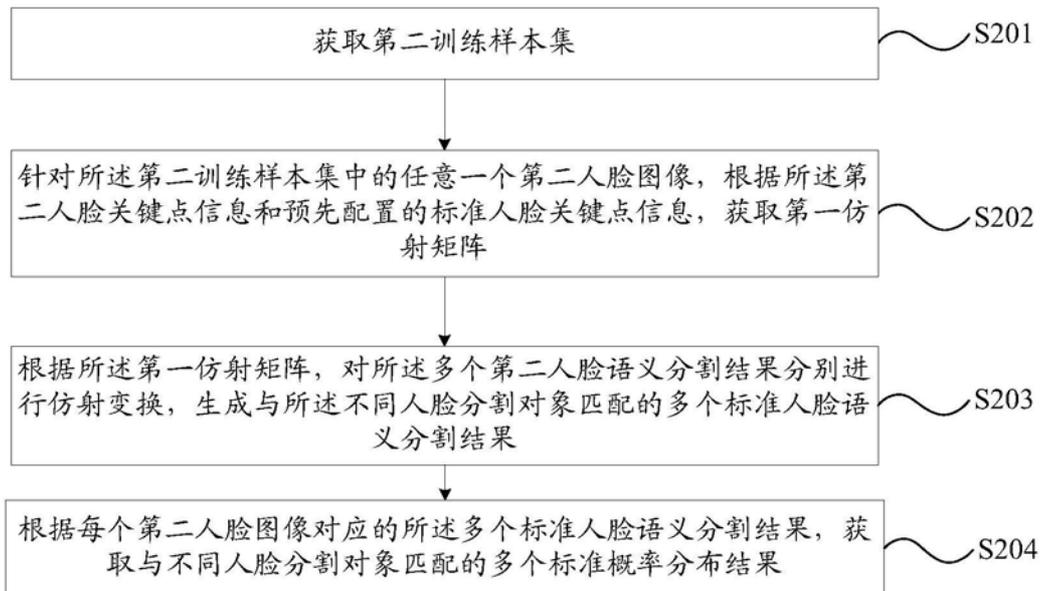


图3

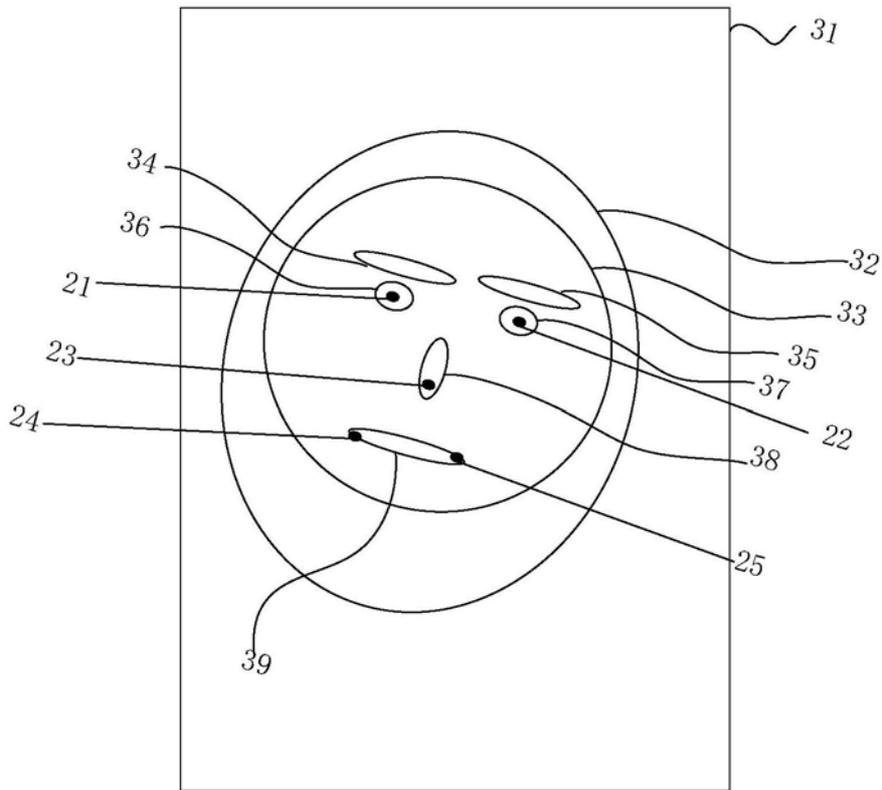


图4

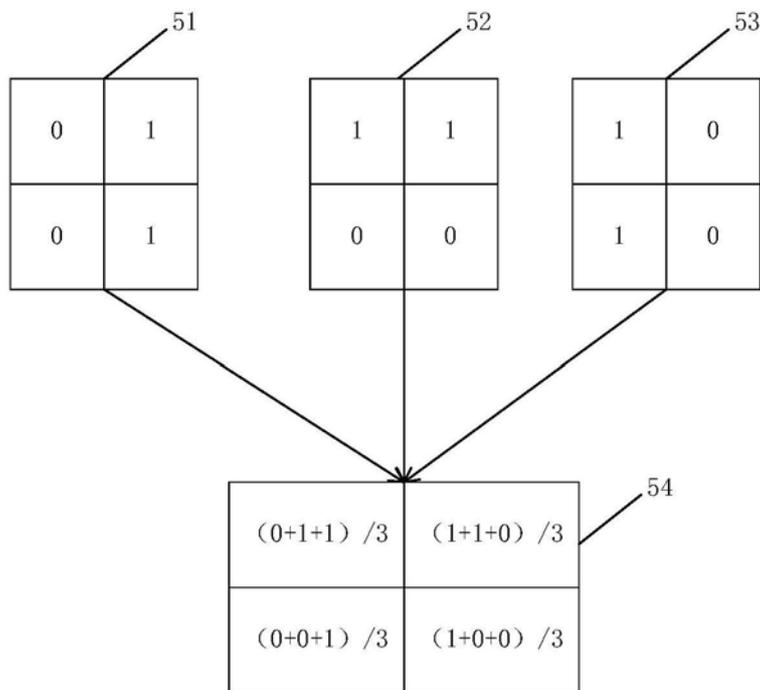


图5

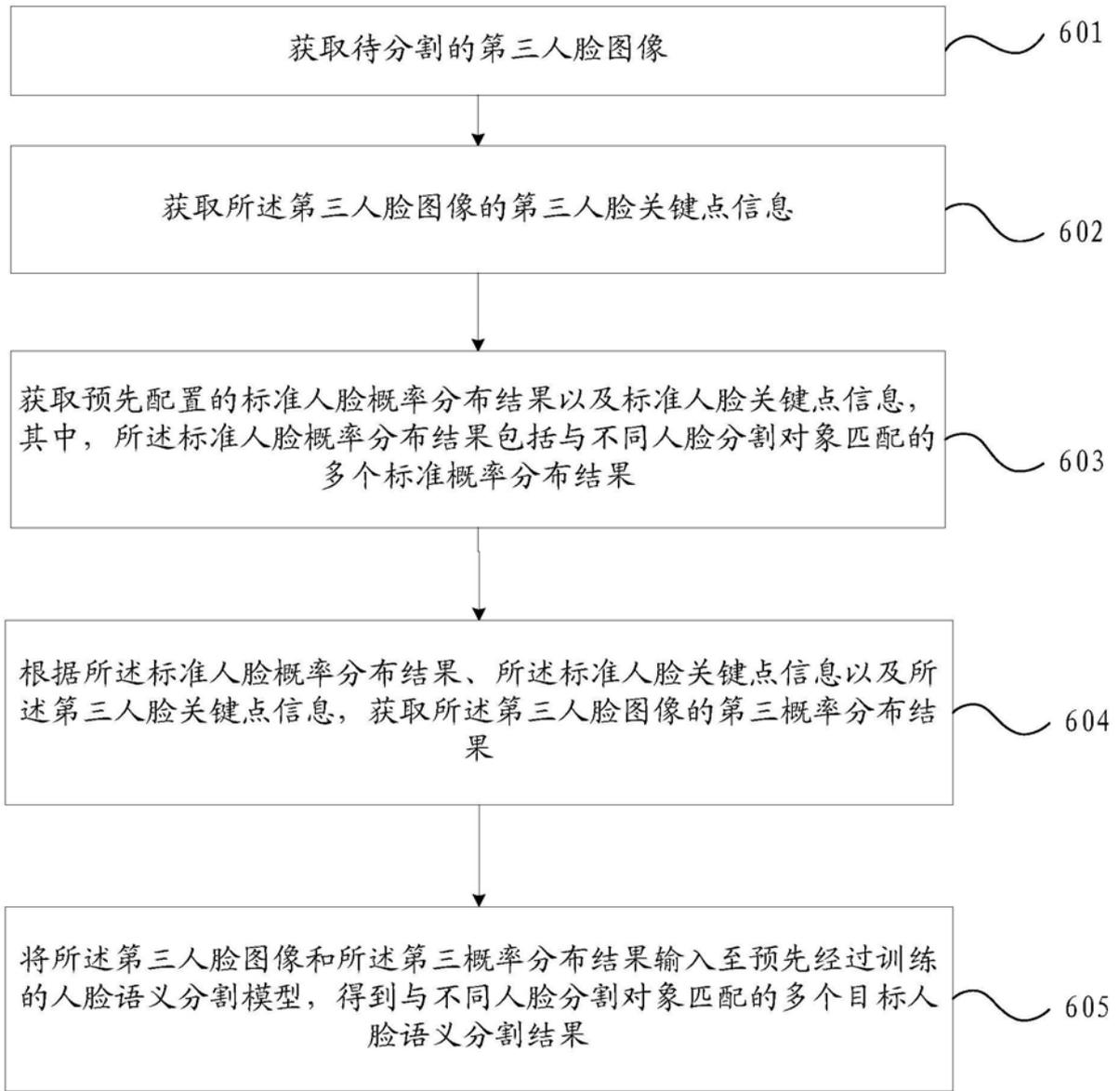


图6

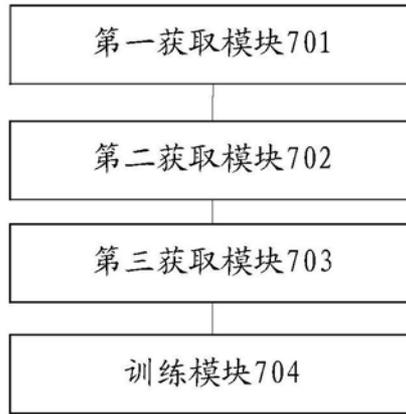


图7

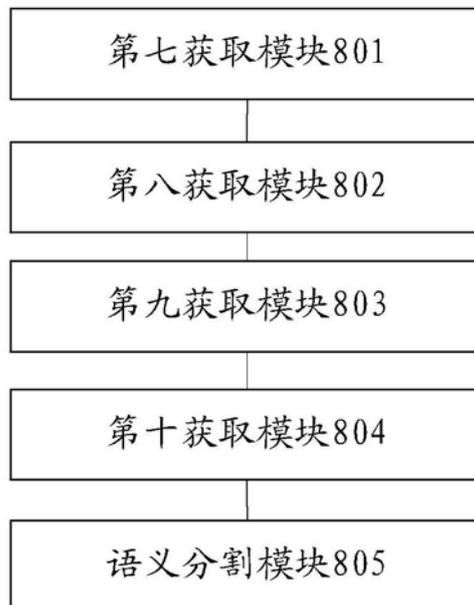


图8