



(12)发明专利申请

(10)申请公布号 CN 107247947 A

(43)申请公布日 2017. 10. 13

(21)申请号 201710552180.8

(22)申请日 2017.07.07

(71)申请人 北京智慧眼科技股份有限公司
地址 100193 北京市海淀区西北旺路10号
院东区14号楼A座4层403

(72)发明人 杨光磊 杨东 王栋

(74)专利代理机构 长沙智嵘专利代理事务所
43211

代理人 胡亮

(51) Int. Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

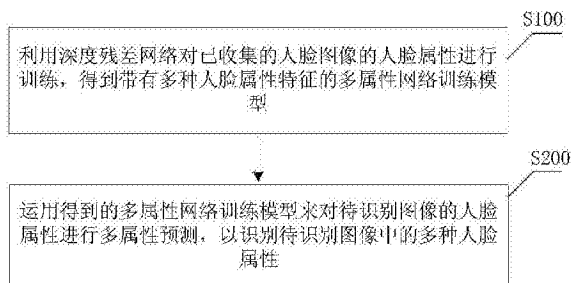
权利要求书2页 说明书11页 附图5页

(54)发明名称

人脸属性识别方法及装置

(57)摘要

本发明公开了一种人脸属性识别方法及装置,利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。本发明提供的人脸属性识别方法及装置,利用一个深度残差网络对已收集的人脸图像的人脸属性进行训练,根据特定任务训练得到的多属性网络训练模型对待识别图像的人脸属性进行多属性预测,因多属性网络模型能提取出更强表达能力的人脸特征,从而可快速准确识别人脸图像中的多种人脸属性。



1. 一种人脸属性识别方法,其特征在于,包括步骤:

利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;

运用得到的所述多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别所述待识别图像中的多种人脸属性。

2. 根据权利要求1所述的人脸属性识别方法,其特征在于,

所述利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型的步骤包括:

将已收集的人脸图像输入深度残差网络,利用所述深度残差网络的网络结构提取所述人脸图像中的人脸属性特征,以在所述深度残差网络中输出人脸特征向量;

对所述深度残差网络进行拓展,将所述深度残差网络输出的所述人脸特征向量集中训练成多属性网络模型。

3. 根据权利要求2所述的人脸属性识别方法,其特征在于,

所述对所述深度残差网络进行拓展,将所述深度残差网络输出的所述多属性特征向量集中训练成多属性网络模型的步骤包括:

在所述深度残差网络的网络结构后面添加深度神经网络的全连接层,使所述全连接层与所述深度残差网络输出的所述人脸特征向量对应连接;

将所述深度残差网络中输出的所述人脸特征向量对应输入到所述全连接层上,得到设定维度的人脸特征训练图像的多属性特征。

4. 根据权利要求3所述的人脸属性识别方法,其特征在于,

所述将所述深度残差网络中输出的所述人脸特征向量对应输入到所述全连接层上,得到设定维度的人脸特征训练图像的多属性特征的步骤之后包括:

在所述深度残差网络的网络结构后面添加深度神经网络的损失函数层,使所述损失函数层与所述深度残差网络中的输出层对应相连,获取所述多属性人脸特征训练图像中的训练数据集标签,计算出所述深度残差网络中所述多属性人脸特征训练图像前向传播时的损失。

5. 根据权利要求1至4中任一项所述的人脸属性识别方法,其特征在于,

所述运用得到的所述多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别所述待识别图像中的多种人脸属性的步骤包括:

将所述待识别图像输入所述多属性网络训练模型,提取所述待识别图像中的多属性人脸特征;

将提取的所述多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征;

将所述全连接层的输出结果送入softmax层,计算出所述多属性人脸特征的最大概率值,以识别所述待识别图像中的多种人脸属性。

6. 一种人脸属性识别装置,其特征在于,包括:

训练模块(10),用于利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;

预测模块(20),用于运用得到的所述多属性网络训练模型来对待识别图像的人脸属性

进行多属性预测,以识别所述待识别图像中的多种人脸属性。

7. 根据权利要求6所述的人脸属性识别装置,其特征在于,

所述训练模块(10)包括第一提取单元(11)和拓展单元(12),

所述第一提取单元(11),用于利用所述深度残差网络的网络结构提取所述人脸图像中的人脸属性特征,以在所述深度残差网络中输出人脸特征向量;

所述拓展单元(12),用于对所述深度残差网络进行拓展,将所述深度残差网络输出的所述多属性特征向量集中训练成多属性网络模型。

8. 根据权利要求7所述的人脸属性识别装置,其特征在于,

所述拓展单元(12)包括连接子单元(121)和训练图像获取子单元(122),

所述连接子单元(121),用于在所述深度残差网络的网络结构后面添加深度神经网络的全连接层,使所述全连接层与所述深度残差网络输出的所述人脸特征向量对应连接;

所述训练图像获取子单元(122),用于将所述深度残差网络中输出的所述人脸特征向量对应输入到所述全连接层上,得到设定维度的人脸特征训练图像的多属性特征。

9. 根据权利要求8所述的人脸属性识别装置,其特征在于,

所述拓展单元(12)还包括损失计算子单元(123),

所述损失计算子单元(123),用于在所述深度残差网络的网络结构后面添加深度神经网络的损失函数层,使所述损失函数层与所述深度残差网络中的输出层对应相连,获取所述多属性人脸特征训练图像中的训练数据集标签,计算出所述深度残差网络中所述多属性人脸特征训练图像前向传播时的损失。

10. 根据权利要求6至9中任一项所述的人脸属性识别装置,其特征在于,

所述预测模块(20)包括第二提取单元(21)、获取单元(22)和计算单元(23),

所述第二提取单元(21),用于将所述待识别图像输入所述多属性网络训练模型,提取所述待识别图像中的多属性人脸特征;

所述获取单元(22),用于将提取的所述多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征;

所述计算单元(23),用于将所述全连接层的输出结果送入softmax层,计算出所述多属性人脸特征的最大概率值,以识别所述待识别图像中的多种人脸属性。

人脸属性识别方法及装置

技术领域

[0001] 本发明涉及人脸识别技术领域,特别地,涉及一种人脸属性识别方法及装置。

背景技术

[0002] 人脸属性识别技术是通过检测人脸图像,获得该人脸图像的年龄、性别等属性信息。在目前的生活中,安全问题成为了大家关注的重心。通过监控视频获取行人的年龄、性别等属性则能够更好的协助司法人员匹配嫌疑人的特征,协助办案。对于购物网站,计算机系统自动分析购物者年龄性别属性后则能够自动向消费者推荐与其身份相符的商品。除此之外,包含人脸图像属性识别技术的应用软件也具有更强的娱乐特性。

[0003] 早期的人脸属性检测技术多为利用传统的SVM (Support Vector Machine,支撑向量机) 进行分类的方法,对于性别、是否戴眼镜等属性可以利用分类器实现转化为二分类问题,对于年龄则将其转化为不同年龄或不同年龄段的多分类问题。现阶段深度学习技术因其卓越的性能而被广泛使用,在人脸属性识别中也有很多应用。

[0004] 然而,基于人脸图像的属性识别很大程度受人脸角度、姿态和光线等因素的影响,识别难度很大,传统的利用SVM的方法很难做到很高的准确率,该方法容易受到环境因素影响。

[0005] Gil Levi和Tal Hassner提出的利用深度学习进行属性识别的方法,识别出年龄的精度低,只能判断出该人脸图像属于哪个年龄段。另外他们使用的深度残差网络模型层数少,对于图像的表达能力弱,所以方法鲁棒性较弱,识别结果的准确性较差。

[0006] 刘昕等人的方法由于使用了8个深层的网络,虽然年龄识别准确度较高,但是识别速度慢,无法做到实时的属性检测。另外他们的方法只实现了年龄属性的识别,而没有做其他属性的识别。

[0007] 因此,现有的人脸属性检测技术在识别类型和识别准确度上存在的缺陷,是一个亟待解决的技术问题。

发明内容

[0008] 本发明提供了一种人脸属性识别方法及装置,以解决现有的人脸属性检测技术在识别类型和识别准确度上存在的缺陷的技术问题。

[0009] 本发明采用的技术方案如下:

[0010] 根据本发明的一个方面,提供一种人脸属性识别方法,包括步骤:

[0011] 利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;

[0012] 运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。

[0013] 进一步地,利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型的步骤包括:

[0014] 将已收集的人脸图像输入深度残差网络,利用深度残差网络的网络结构提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量;

[0015] 对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型。

[0016] 进一步地,对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型的步骤包括:

[0017] 在深度残差网络的网络结构后面添加深度神经网络的全连接层,使全连接层与深度残差网络输出的人脸特征向量对应连接;

[0018] 将深度残差网络中输出的人脸特征向量对应输入到全连接层上,得到设定维度的人脸特征训练图像的多属性特征。

[0019] 进一步地,将深度残差网络中输出的人脸特征向量对应输入到全连接层上,得到设定维度的人脸特征训练图像的多属性特征的步骤之后包括:

[0020] 在深度残差网络的网络结构后面添加深度神经网络的损失函数层,使损失函数层与深度残差网络中的输出层对应相连,获取多属性人脸特征训练图像中的训练数据集标签,计算出深度残差网络中多属性人脸特征训练图像前向传播时的损失。

[0021] 进一步地,运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性的步骤包括:

[0022] 将待识别图像输入多属性网络训练模型,提取待识别图像中的多属性人脸特征;

[0023] 将提取的多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征;

[0024] 将全连接层的输出结果送入softmax层,计算出多属性人脸特征的最大概率值,以识别待识别图像中的多种人脸属性。

[0025] 根据本发明的另一方面,还提供一种人脸属性识别装置,包括:

[0026] 训练模块,用于利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;

[0027] 预测模块,用于运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。

[0028] 进一步地,训练模块包括第一提取单元和拓展单元,

[0029] 第一提取单元,用于将已收集的人脸图像输入深度残差网络,利用深度残差网络的网络结构提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量;

[0030] 拓展单元,用于对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型。

[0031] 进一步地,拓展单元包括连接子单元和训练图像获取子单元,

[0032] 连接子单元,用于在深度残差网络的网络结构后面添加深度神经网络的全连接层,使全连接层与深度残差网络输出的人脸特征向量对应连接;

[0033] 训练图像获取子单元,用于将深度残差网络中输出的人脸特征向量对应输入到全连接层上,得到设定维度的人脸特征训练图像的多属性特征。

[0034] 进一步地,拓展单元还包括损失计算子单元,

[0035] 损失计算子单元,用于在深度残差网络的网络结构后面添加深度神经网络的损失

函数层,使损失函数层与深度残差网络中的输出层对应相连,获取多属性人脸特征训练图像中的训练数据集标签,计算出深度残差网络中多属性人脸特征训练图像前向传播时的损失。

[0036] 进一步地,预测模块包括第二提取单元、获取单元和计算单元,

[0037] 第二提取单元,用于将待识别图像输入多属性网络训练模型,提取待识别图像中的多属性人脸特征;

[0038] 获取单元,用于将提取的多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征;

[0039] 计算单元,用于将全连接层的输出结果送入softmax层,计算出多属性人脸特征的最大概率值,以识别待识别图像中的多种人脸属性。

[0040] 本发明具有以下有益效果:

[0041] 本发明提供的人脸属性识别方法及装置,通过利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。本发明提供的人脸属性识别方法及装置,利用一个深度残差网络对已收集的人脸图像的人脸属性进行训练,根据特定任务训练得到的多属性网络训练模型对待识别图像的人脸属性进行多属性预测,因多属性网络模型能提取出更强表达能力的人脸特征,从而可快速准确识别人脸图像中的多种人脸属性。

[0042] 除了上面所描述的目的、特征和优点之外,本发明还有其它的目的、特征和优点。下面将参照图,对本发明作进一步详细的说明。

附图说明

[0043] 构成本申请的一部分的附图用来提供对本发明的进一步理解,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

[0044] 图1是本发明人脸属性识别方法第一优选实施例的流程示意图;

[0045] 图2为图1中利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型的步骤优选实施例的细化流程示意图;

[0046] 图3为本发明人脸属性识别方法优选实施例中人脸多属性网络训练流程图;

[0047] 图4为图2中对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型的步骤优选实施例的细化流程示意图;

[0048] 图5为图1中运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性的步骤优选实施例的细化流程示意图;

[0049] 图6为本发明人脸属性识别方法优选实施例中人脸多属性预测流程图;

[0050] 图7为本发明人脸属性识别装置第一优选实施例的功能框图;

[0051] 图8为图7中训练模块优选实施例的功能模块示意图;

[0052] 图9为图8中拓展单元优选实施例的功能模块示意图;

[0053] 图10为图7中预测模块优选实施例的功能模块示意图。

[0054] 附图标号说明:

[0055] 10、训练模块;20、预测模块;11、第一提取单元;12、拓展单元;121、连接子单元;

122、训练图像获取子单元;123、损失计算子单元;21、第二提取单元;22、获取单元;23、计算单元。

具体实施方式

[0056] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本发明。

[0057] 参照图1,本发明的优选实施例提供一种人脸属性识别方法,包括步骤:

[0058] 步骤S100、利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型。

[0059] 由于不同的人脸属性相互间具有相关性,通过一个深度残差网络训练已收集的人脸图像的所有人脸属性,得到带有多种人脸属性特征的多属性网络训练模型。在机器学习中,训练是用带有标签的人脸图像来训练一个机器学习的模型。由于深度残差网络具有更深的网络层数,对于不同的任务训练都取得了很好的效果。其中,人脸属性包括年龄、性别、眼镜、头部姿态、胡须、眼睛状态和微笑程度等中的二种或二种以上。

[0060] 步骤S200、运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。

[0061] 将待识别图像输入到训练得到的多属性网络训练模型中,运用多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性,并输出识别结果。预测是指用训练得到的多属性网络训练模型对未知标签的待识别图像做预测,预测待识别图像的未知标签。在本实施例中,对于性别、眼镜、微笑这种较为容易的二分类问题,直接用Softmax层得出该属性是或否的概率,把最大概率值作为该属性的预测。对于年龄,Softmax层输出与年龄对应的80维年龄预测概率分布,用概率分布与标准年龄基求相似性,用相似性得到的权重对标准年龄基加权求和,最大值对应维度为年龄属性的预测。

[0062] 本实施例提供的人脸属性识别方法,利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。本实施例提供的人脸属性识别方法,利用一个深度残差网络对已收集的人脸图像的人脸属性进行训练,根据特定任务训练得到的多属性网络训练模型对待识别图像的人脸属性进行多属性预测,因多属性网络模型能提取出更强表达能力的人脸特征,从而可快速准确识别人脸图像中的多种人脸属性。

[0063] 优选地,如图2所示,本实施例提供的人脸属性识别方法,步骤S100包括:

[0064] 步骤S110、将已收集的人脸图像输入深度残差网络,利用深度残差网络的网络结构提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量。

[0065] 深度残差网络是在2015年由微软亚洲实验室提出的深度学习网络结构,该结构具有更深的网络层数,在多项图像相关竞赛中夺冠。在本实施例中,利用深度残差网络的网络结构提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量,人脸属性包括年龄、性别、眼镜、头部姿态、胡须、眼睛状态和微笑程度等中的二种或二种以上。

[0066] 步骤S120、对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型。

[0067] 利用网上公布的在ImageNet数据集上训练的网络模型,对提取的人脸图像中的人脸属性特征进行再训练来得到多属性网络模型。原深度残差网络输出为1000维向量,对图像做分类,能够分类1000类物体,对应输出1000维度,每一维度对应识别为某种分类物体的概率。在本实施例中,为了训练多属性网络,对深度残差网络进行拓展,在深度残差网络的网络结构后面添加与每个属性相对应的损失函数层,使每个损失函数层都与深度残差网络的输出层相连。如图3所示,对于残差网络输出的1000维特征向量分别连接不同的全连接层来识别不同任务。

[0068] 本实施例提供的人脸属性识别方法,将已收集的人脸图像输入深度残差网络,利用深度残差网络的网络结构提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量;对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型。本实施例提供的人脸属性识别方法,利用一个深度残差网络实现人脸图像多属性的识别,根据特定任务训练出来的多属性网络模型提取出更强表达能力的人脸特征,实现人脸属性的快速准确的识别。

[0069] 优选地,如图4所示,本实施例提供的人脸属性识别方法,步骤S120包括:

[0070] 步骤S121、在深度残差网络的网络结构后面添加深度神经网络的全连接层,使全连接层与深度残差网络输出的人脸特征向量对应连接。

[0071] 请见图3,由于脸部特征具有相关性,通过在深度残差网络的网络结构后面添加不同的深度神经网络的全连接层,使全连接层与深度残差网络输出的1000维多属性特征向量对应连接。从而使不同属性的识别可以共享同一个神经网络,实现多个属性的同时训练与识别,最大限度加快训练与预测速度。

[0072] 步骤S122、将深度残差网络中输出的人脸特征向量对应输入到全连接层上,得到设定维度的人脸特征训练图像的多属性特征。

[0073] 参见图3,对于深度残差网络输出的1000维特征向量分别连接不同的全连接层来识别不同任务。其中,对于年龄属性,在对应的全连接层 f_{c1} 上输出80维的人脸特征训练图像的年龄属性。对于性别属性,在对应的全连接层 f_{c2} 上输出2维的人脸特征训练图像的性别属性特征。对于眼镜属性,在对应的全连接层 f_{c3} 上输出2维的人脸特征训练图像的眼镜属性特征。

[0074] 步骤S123、在深度残差网络的网络结构后面添加深度神经网络的损失函数层,使损失函数层与深度残差网络中的输出层对应相连,获取多属性人脸特征训练图像中的训练数据集标签,计算出深度残差网络中多属性人脸特征训练图像前向传播时的损失。

[0075] 对于年龄属性,不再使用常规的多年龄的分类方法来训练,而是将待识别图像的年龄标签看作高斯函数的形式,利用交叉熵来计算深度残差网络中多属性人脸特征训练图像前向传播时的损失。对于年龄为age的训练图像,多属性人脸特征训练图像的年龄标签为:

$$[0076] \quad \text{label}_i = \frac{1}{\sqrt{2\pi}\sigma_{\text{age}}} e^{-\frac{(i-\text{age})^2}{\sigma_{\text{age}}^2}} \quad i = 1, \dots, 80 \quad (1)$$

[0077] 其中, σ_{age} 为该年龄对应标准差, i 取值从1到80表示预测年龄范围为1岁到80岁。对于不同年龄设置 σ_{age} 有区别,在设定的年龄阈值范围外时,例如较小年龄(<10岁)以及较大

年龄(>70岁)设置更小的 α_{age} ,从而可以避免边缘年龄识别向中间年龄段偏移的问题。

[0078] 将全连接层 f_{c1} 输出的多属性人脸特征训练图像中的80维年龄特征通过softmax层,输出多属性人脸特征训练图像属于每个年龄的概率,softmax运算如下:

$$[0079] \quad p_i = \frac{e^{x_i}}{\sum_{t=1}^{80} e^{x_t}} \quad (2)$$

[0080] 其中, x 为输入向量, p 为输出概率。交叉熵函数能够更好的度量两个概率分布之间的距离,因此计算输出年龄概率 $p = [p_1, p_2, \dots, p_{80}]$ 和标签 $q = [q_1, q_2, \dots, q_{80}]$ 之间的损失使用的交叉熵代价函数,对于训练样本标签,如下式所示:

$$[0081] \quad \text{loss_crossentropy} = -\frac{1}{N} \sum_{t=1}^N \sum_{i=1}^{80} q_t^i \log(p_t^i) \quad (3)$$

[0082] 其中, N 为训练图像张数, i 为对应维度。

[0083] 对于除年龄属性外的其他属性使用Softmax损失函数计算代价并进行优化,Softmax代价函数如下所示:

$$[0084] \quad \text{loss_softmax} = -\frac{1}{N} \sum_{i=1}^N p_j \quad (4)$$

[0085] 其中, p_j 为softmax运算概率值, j 为多属性人脸特征训练图像真实的类别编号。

[0086] 通过以上运算,可以计算每一个属性对应的代价。并通过反向传播算法更新网络参数。

[0087] 本实施例提供的人脸属性识别方法,通过将深度残差网络中输出的人脸特征向量对应输入到全连接层上,得到设定维度的人脸特征训练图像的多属性特征;在深度残差网络的网络结构后面添加深度神经网络的损失函数层,使损失函数层与深度残差网络中的输出层对应相连,获取多属性人脸特征训练图像中的训练数据集标签,计算出深度残差网络中多属性人脸特征训练图像前向传播时的损失。本实施例提供的人脸属性识别方法,通过添加不同的神经网络的全连接层,从而使不同属性的识别共享同一个神经网络,实现多个属性的同时训练与识别,最大限度加快训练与预测速度;对于所有属性中难度较大的年龄识别问题,将所有年龄图像的标签表示为有方差区别的高斯函数的形式,采用交叉熵作为代价函数做训练,以得出合理的年龄预测方案,该方案通过一种软分配的方法使得年龄识别具有更强的鲁棒性。

[0088] 优选地,如图5所示,本实施例提供的人脸属性识别方法,步骤S200包括:

[0089] 步骤S210、将待识别图像输入多属性网络训练模型,提取待识别图像中的多属性人脸特征。

[0090] 人脸多属性预测的整体方案如图6所示。将待识别图像输入多属性网络训练模型,通过多属性网络训练模型提取待识别图像中的多属性人脸特征,网络输出1000维特征向量。

[0091] 步骤S220、将提取的多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征。

[0092] 对于年龄属性,将多属性网络训练模型网络输出的1000维特征向量输入对应的全连接层 f_{c1} ,在全连接层 f_{c1} 中输出具有80维年龄特征的人脸图像多属性预测特征。对于性

别、是否戴眼镜等其他属性,将1000维特征向量输入对应的全连接层 f_{c_i} ,在全连接层 f_{c_i} 中输出具有2维特征的人脸图像多属性预测特征。

[0093] 步骤S230、将全连接层的输出结果送入softmax层,计算出多属性人脸特征的最大概率值,以识别待识别图像中的多种人脸属性。

[0094] 对于年龄属性,将全连接层 f_{c_i} 输出结果送入对应的Softmax层,以计算出初始的各年龄预测概率的结果,输出预测年龄。对于性别、是否戴眼镜等其他属性,通过对应的Softmax层计算出2维特征属性是与否的概率,并取最大概率对应情况判断2维特征属性情况。

[0095] 在本实施例中,网络训练结束后通过多属性网络训练模型预测待识别图像的多种人脸属性。对于除年龄外的其他属性,参考图6可得到。对于年龄属性,构造一组标准特征基。因为网络输出特征结果可能存在误差,但相似年龄图像输出特征是相似的,通过特征比对找到相似年龄图像并利用软分配的方法将年龄标签加权求和,实现更准确的年龄识别。

[0096] 对于每个年龄挑选N张图像作为标准年龄图像,生成维度与年龄对应的80维年龄特征 f (即图6中 f_{c_i} 通过Softmax运算后的结果)作为标准年龄特征基。对于每张待识别图像同样提取80维年龄特征 f ,并计算80维年龄特征 f 与标准年龄特征基之间的权重 ε_i :

$$[0097] \quad \varepsilon_i = e^{-\|f-f_i^{set}\|_2^2} \quad (5)$$

[0098] 其中, f 为80维年龄特征, f_i^{set} 为特征基中第 i 幅图像特征。

[0099] 将 ε_i 乘以图像 i 对应的标准的标签label作为图像 i 对于年龄预测的贡献值,将年龄特征基中所有图像贡献值对应维度相加所得到的80维向量,则为最终预测结果,80维向量中最大值对应维度即为预测年龄。

[0100] 本实施例提供的人脸属性识别方法,将待识别图像输入多属性网络训练模型,提取待识别图像中的多属性人脸特征;将提取的多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征;将全连接层的输出结果送入softmax层,计算出多属性人脸特征的最大概率值,以识别待识别图像中的多种人脸属性。本实施例提供的人脸属性识别方法,由于脸部特征都具有相关性,通过添加不同的神经网络全连接层,不同属性的识别可以共享同一个神经网络,实现多个属性的同时训练与识别,最大限度加快训练与预测速度;预测时利用年龄特征基里的标准标签加权求和来选出年龄预测结果。对于训练好的多属性网络训练模型,生成与各个年龄段对应维度的年龄特征,并将生成的对应维度的年龄特征作为年龄特征基。在待识别图像中提取年龄特征向量,并将提取的年龄特征向量与所有年龄特征基计算距离,用距离的指数函数加权年龄特征基所对应的待识别图像的标准标签,将所有加权后的标准标签相加后求和,最大值对应维度则为年龄预测值。该方案通过一种软分配的方法使得年龄识别具有更强的鲁棒性。

[0101] 优选地,如图7所示,本实施例还提供一种人脸属性识别装置,包括:训练模块10,用于利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;预测模块20,用于运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。

[0102] 由于不同的人脸属性相互间具有相关性,通过训练模块10训练已收集的人脸图像的所有人脸属性,得到带有多种人脸属性特征的多属性网络训练模型。在机器学习中,训练

是用带有标签的人脸图像来训练一个机器学习的模型。由于深度残差网络具有更深的网络层数,对于不同的任务训练都取得了很好的效果。其中,人脸属性包括年龄、性别、眼镜、头部姿态、胡须、眼睛状态和微笑程度等中的二种或二种以上。

[0103] 将待识别图像输入到训练得到的多属性网络训练模型中,预测模块20运用多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性,并输出识别结果。预测是指用训练得到的多属性网络训练模型对未知标签的待识别图像做预测,预测待识别图像的未知标签。在本实施例中,对于性别、眼镜、微笑这种较为容易的二分类问题,直接用Softmax层得出该属性是或否的概率,把最大概率值作为该属性的预测。对于年龄,Softmax层输出与年龄对应的80维年龄预测概率分布,用概率分布与标准年龄基求相似性,用相似性得到的权重对标准年龄基加权求和,最大值对应维度为年龄属性的预测。

[0104] 本实施例提供的人脸属性识别装置,通过利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。本实施例提供的人脸属性识别装置,通过训练得到的多属性网络训练模型对待识别图像的人脸属性进行多属性预测,可快速准确识别人脸图像中的多种人脸属性。利用深度残差网络对已收集的人脸图像的人脸属性进行训练,得到带有多种人脸属性特征的多属性网络训练模型;运用得到的多属性网络训练模型来对待识别图像的人脸属性进行多属性预测,以识别待识别图像中的多种人脸属性。本实施例提供的人脸属性识别方法,利用一个深度残差网络对已收集的人脸图像的人脸属性进行训练,根据特定任务训练得到的多属性网络训练模型对待识别图像的人脸属性进行多属性预测,因多属性网络模型能提取出更强表达能力的人脸特征,从而可快速准确识别人脸图像中的多种人脸属性。

[0105] 优选地,本实施例提供的人脸属性识别装置,训练模块10包括第一提取单元11和拓展单元12,第一提取单元11,用于将已收集的人脸图像输入深度残差网络,利用深度残差网络的网络结构提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量;拓展单元12,用于对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型。

[0106] 深度残差网络是在2015年由微软亚洲实验室提出的深度学习网络结构,该结构具有更深的网络层数,在多项图像相关竞赛中夺冠。在本实施例中,利用第一提取单元11提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量,人脸属性包括年龄、性别、眼镜、头部姿态、胡须、眼睛状态和微笑程度等中的二种或二种以上。

[0107] 利用网上公布的在ImageNet数据集上训练的网络模型,对提取的人脸图像中的人脸属性特征进行再训练来得到网络模型。原深度残差网络输出为1000维特征向量,对图像做分类,能够分类1000类物体,对应输出1000维度,每一维度对应识别为某种分类物体的概率。在本实施例中,为了训练多属性网络,拓展单元12在深度残差网络进行拓展,在深度残差网络的网络结构后面添加每个属性的损失函数层,使每个损失函数层都与原深度残差网络的输出层相连。如图3所示,对于残差网络输出的1000维特征向量分别连接不同的全连接层来识别不同任务。

[0108] 本实施例提供的人脸属性识别装置,将已收集的人脸图像输入深度残差网络,利

用深度残差网络的网络结构提取人脸图像中的人脸属性特征,以在深度残差网络中输出人脸特征向量;对深度残差网络进行拓展,将深度残差网络输出的人脸特征向量集中训练成多属性网络模型。本实施例提供的人脸属性识别装置,利用一个深度残差网络实现人脸图像多属性的识别,根据特定任务训练出来的多属性网络模型提取出更强表达能力的特征,实现人脸属性的快速准确的识别。

[0109] 优选地,本实施例提供的人脸属性识别装置,拓展单元12包括连接子单元121、训练图像获取子单元122和损失计算子单元123,连接子单元121,用于在深度残差网络的网络结构后面添加深度神经网络的全连接层,使全连接层与深度残差网络输出的人脸特征向量对应连接。训练图像获取子单元122,用于将深度残差网络中输出的人脸特征向量对应输入到全连接层上,得到设定维度的人脸特征训练图像的多属性特征。损失计算子单元123,用于在深度残差网络的网络结构后面添加深度神经网络的损失函数层,使损失函数层与深度残差网络中的输出层对应相连,获取多属性人脸特征训练图像中的训练数据集标签,计算出深度残差网络中多属性人脸特征训练图像前向传播时的损失。

[0110] 请见图3,由于脸部特征具有相关性,通过在深度残差网络的网络结构后面添加不同的深度神经网络的全连接层,使全连接层与深度残差网络输出的1000维多属性特征向量对应连接。从而使不同属性的识别可以共享同一个神经网络,实现多个属性的同时训练与识别,最大限度加快训练与预测速度。

[0111] 参见图3,对于深度残差网络输出的1000维特征向量分别连接不同的全连接层来识别不同任务。其中,对于年龄属性,在对应的全连接层 f_{c1} 上输出80维的人脸特征训练图像的年龄属性。对于性别属性,在对应的全连接层 f_{c2} 上输出2维的人脸特征训练图像的性别属性特征。对于眼镜属性,在对应的全连接层 f_{c3} 上输出2维的人脸特征训练图像的眼镜属性特征。

[0112] 对于年龄属性,不再使用常规的多年龄的分类方法来训练,而是通过损失计算子单元123将待识别图像的年龄标签看作高斯函数的形式,利用交叉熵来计算深度残差网络中多属性人脸特征训练图像前向传播时的损失。对于年龄为age的训练图像,多属性人脸特征训练图像的年龄标签为:

$$[0113] \quad \text{label}_i = \frac{1}{\sqrt{2\pi}\sigma_{\text{age}}} e^{-\frac{(i-\text{age})^2}{\sigma_{\text{age}}^2}} \quad i=1, \dots, 80 \quad (6)$$

[0114] 其中, σ_{age} 为该年龄对应标准差, i 取值从1到80表示预测年龄范围为1岁到80岁。对于不同年龄设置 σ_{age} 有区别,在设定的年龄阈值范围外时,例如较小年龄(<10岁)以及较大年龄(>70岁)设置更小的 σ_{age} ,从而可以避免边缘年龄识别向中间年龄段偏移的问题。

[0115] 将全连接层 f_{c1} 输出的多属性人脸特征训练图像中的80维年龄特征通过softmax层,输出多属性人脸特征训练图像属于每个年龄的概率,softmax运算如下:

$$[0116] \quad p_i = \frac{e^{x_i}}{\sum_{t=1}^{80} e^{x_t}} \quad (7)$$

[0117] 其中, x 为输入向量, p 为输出概率。交叉熵函数能够更好的度量两个概率分布之间的距离,因此计算输出年龄概率 $p = [p_1, p_2, \dots, p_{80}]$ 和标签 $q = [q_1, q_2, \dots, q_{80}]$ 之间的损失使用的交叉熵代价函数,对于训练样本标签,如下式所示:

$$[0118] \quad \text{loss_crossentropy} = -\frac{1}{N} \sum_{i=1}^N \sum_{i=1}^{80} q_i \log(p_i) \quad (8)$$

[0119] 其中,N为训练图像张数,i为对应维度。

[0120] 对于除年龄属性外的其他属性使用Softmax损失函数计算代价并进行优化,Softmax代价函数如下所示:

$$[0121] \quad \text{loss_softmax} = -\frac{1}{N} \sum_{i=1}^N p_j \quad (9)$$

[0122] 其中, p_j 为softmax运算概率值,j为多属性人脸特征训练图像真实的类别编号。

[0123] 通过以上运算,可以计算每一个属性对应的代价。并通过反向传播算法更新网络参数。

[0124] 本实施例提供的人脸属性识别装置,通过将深度残差网络中输出的人脸特征向量对应输入到全连接层上,得到设定维度的人脸特征训练图像的多属性特征;在深度残差网络的神经网络后面添加深度神经网络的损失函数层,使损失函数层与深度残差网络中的输出层对应相连,获取多属性人脸特征训练图像中的训练数据集标签,计算出深度残差网络中多属性人脸特征训练图像前向传播时的损失。本实施例提供的人脸属性识别方法装置,通过添加不同的神经网络的全连接层,从而使不同属性的识别共享同一个神经网络,实现多个属性的同时训练与识别,最大限度加快训练与预测速度;对于所有属性中难度较大的年龄识别问题,将所有年龄图像的标签表示为有方差区别的高斯函数的形式,采用交叉熵作为代价函数做训练,以得出合理的年龄预测方案,该方案通过一种软分配的方法使得年龄识别具有更强的鲁棒性。

[0125] 优选地,本实施例提供的人脸属性识别装置,预测模块20包括第二提取单元21、获取单元22和计算单元23,第二提取单元21,用于将待识别图像输入多属性网络训练模型,提取待识别图像中的多属性人脸特征;获取单元22,用于将提取的多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征;计算单元23,用于将全连接层的输出结果送入softmax层,计算出多属性人脸特征的最大概率值,以识别待识别图像中的多种人脸属性。

[0126] 人脸多属性预测的整体方案如图6所示。将待识别图像输入多属性网络训练模型,第二提取单元21通过多属性网络训练模型提取待识别图像中的多属性人脸特征,网络输出1000维特征向量。

[0127] 对于年龄属性,获取单元22将多属性网络训练模型网络输出的1000维特征向量输入对应的全连接层 f_{c1} ,在全连接层 f_{c1} 中输出具有80维年龄特征的人脸图像多属性预测特征。对于性别、是否戴眼镜等其他属性,获取单元22将1000维特征向量输入对应的全连接层 f_{ci} ,在全连接层 f_{ci} 中输出具有2维特征的人脸图像多属性预测特征。

[0128] 对于年龄属性,计算单元23将全连接层 f_{c1} 输出结果送入对应的Softmax层,以计算出初始的各年龄预测概率的结果,输出预测年龄。对于性别、是否戴眼镜等其他属性,计算单元23通过对应的Softmax层计算出2维特征属性是与否的概率,并取最大概率对应情况判断2维特征属性情况。

[0129] 在本实施例中,网络训练结束后通过多属性网络训练模型预测待识别图像的多种人脸属性。对于除年龄外的其他属性,参考图6可得到。对于年龄属性,构造一组标准特征

基。因为网络输出特征结果可能存在误差,但相似年龄图像输出特征是相似的,通过特征比对找到相似年龄图像并利用软分配的方法将年龄标签加权求和,实现更准确的年龄识别。

[0130] 对于每个年龄挑选N张图像作为标准年龄图像,生成维度与年龄对应的80维年龄特征f(即图6中 f_{c_i} 通过Softmax运算后的结果)作为标准年龄特征基。对于每张待识别图像同样提取80维年龄特征f,并计算80维年龄特征f与标准年龄特征基之间的权重 ε_i :

$$[0131] \quad \varepsilon_i = e^{-\|f - f_i^{\text{set}}\|_2^2} \quad (10)$$

[0132] 其中,f为80维年龄特征, f_i^{set} 为特征基中第i幅图像特征。

[0133] 将 ε_i 乘以图像i对应的标准的标签label作为图像i对于年龄预测的贡献值,将年龄特征基中所有图像贡献值对应维度相加所得到的80维向量,则为最终预测结果,80维向量中最大值对应维度即为预测年龄。

[0134] 本实施例提供的人脸属性识别装置,将待识别图像输入多属性网络训练模型,提取待识别图像中的多属性人脸特征;将提取的多属性人脸特征分别输入对应的全连接层,得到设定维度的人脸图像多属性预测特征;将全连接层的输出结果送入softmax层,计算出多属性人脸特征的最大概率值,以识别待识别图像中的多种人脸属性。本实施例提供的人脸属性识别装置,由于脸部特征都具有相关性,通过添加不同的神经网络全连接层,不同属性的识别可以共享同一个神经网络,实现多个属性的同时训练与识别,最大限度加快训练与预测速度;预测时利用年龄特征基里的标准标签加权求和来选出年龄预测结果。对于训练好的多属性网络训练模型,生成与各个年龄段对应维度的年龄特征,并将生成的对应维度的年龄特征作为年龄特征基。在待识别图像中提取年龄特征向量,并将提取的年龄特征向量与所有年龄特征基计算距离,用距离的指数函数加权年龄特征基所对应的待识别图像的标准标签,将所有加权后的标准标签相加后求和,最大值对应维度则为年龄预测值。该方案通过一种软分配的方法使得年龄识别具有更强的鲁棒性。

[0135] 以上所述仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技术人员来说,本发明可以有各种更改和变化。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。

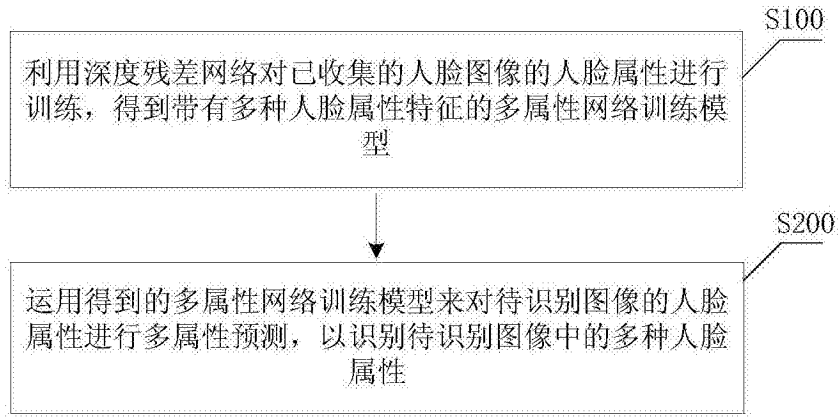


图1

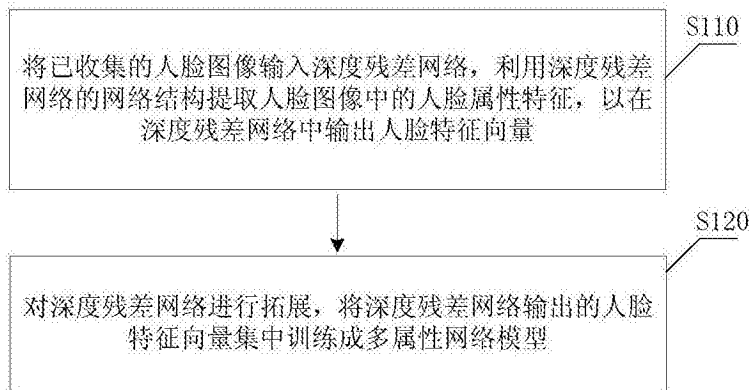


图2

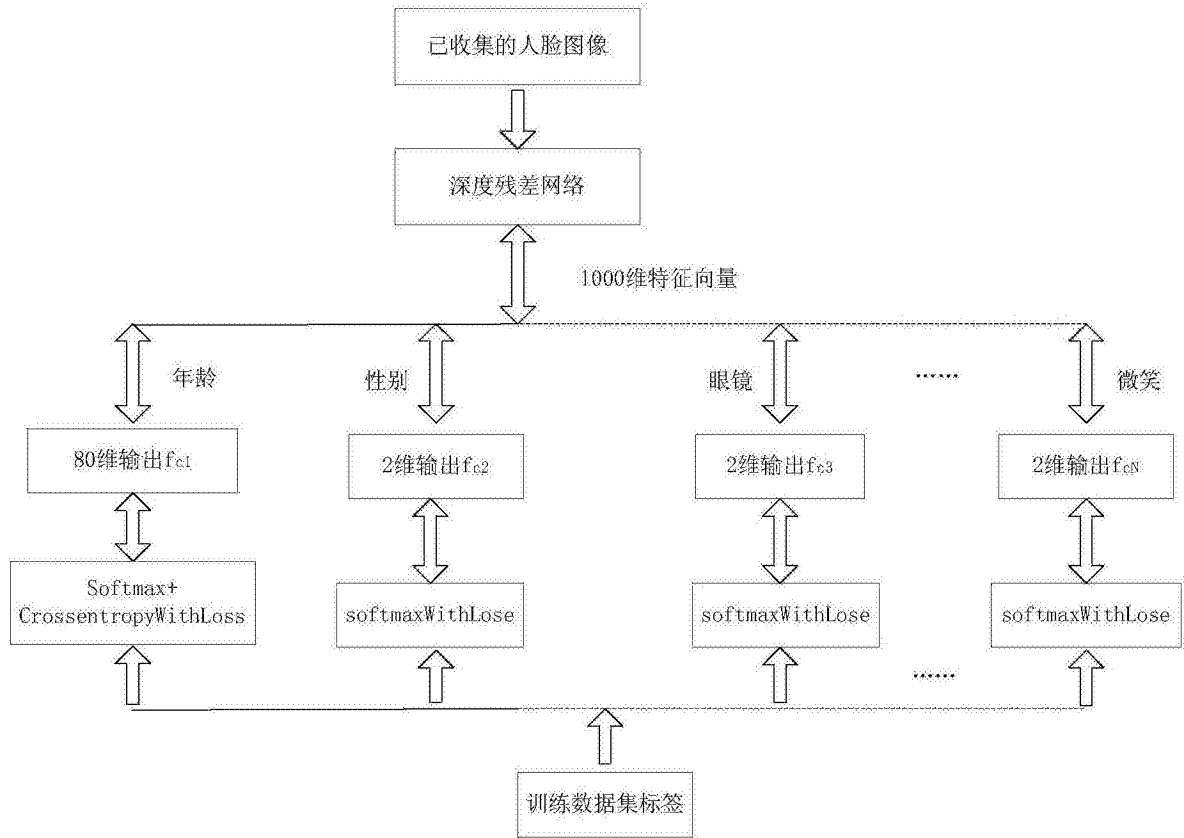


图3

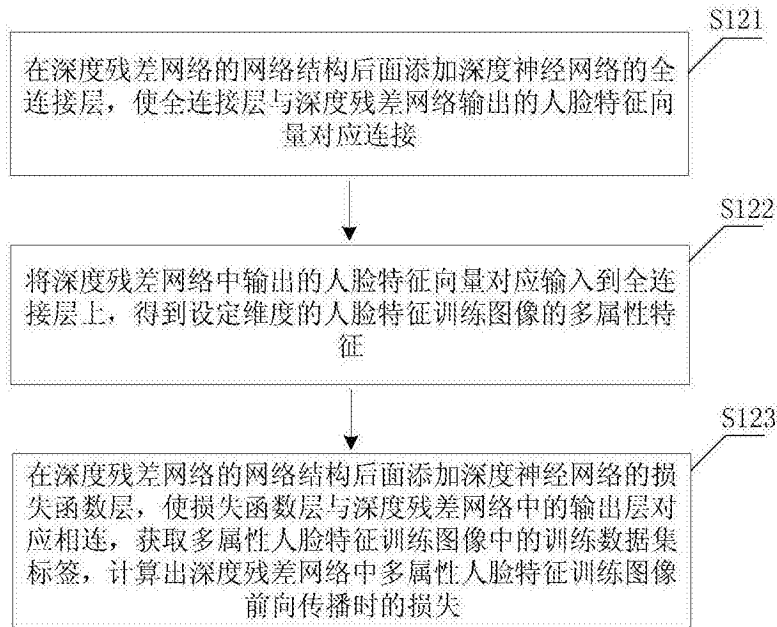


图4

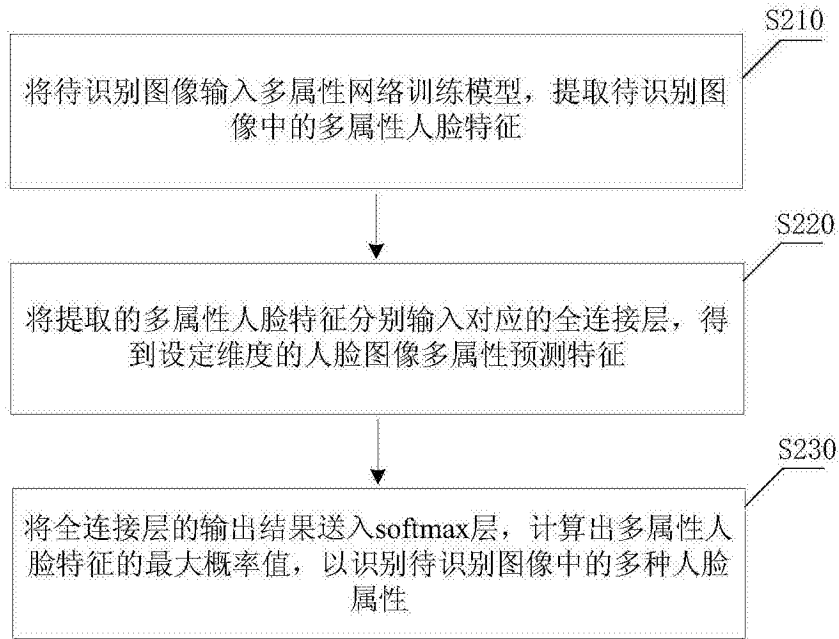


图5

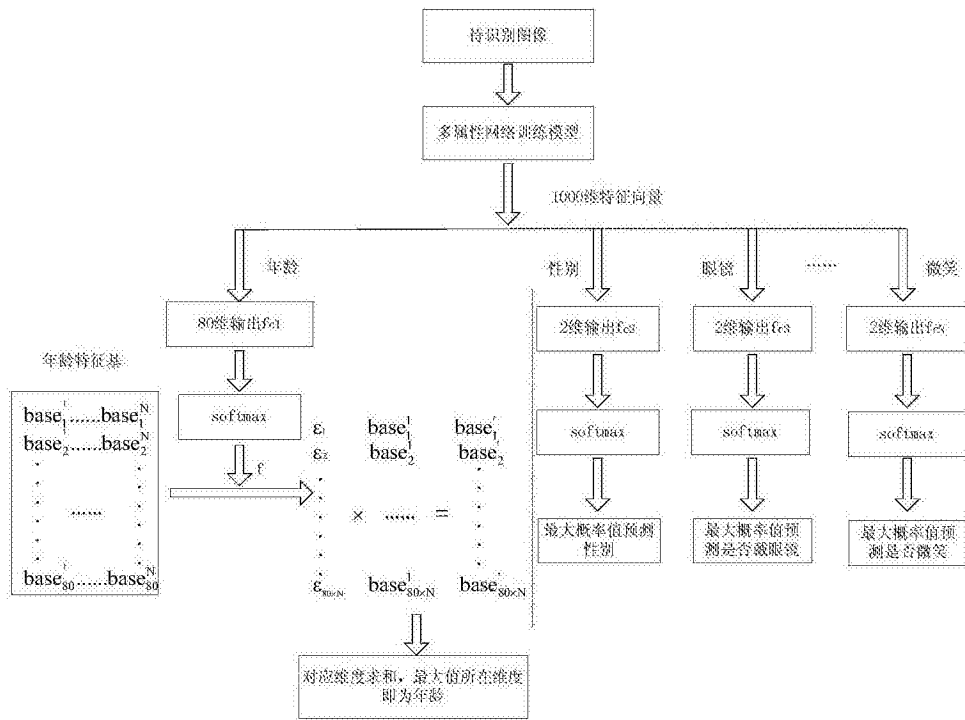


图6

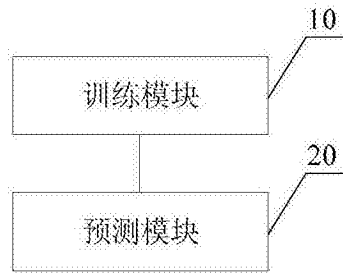


图7

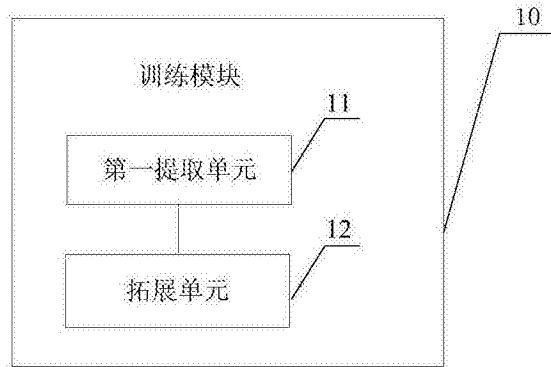


图8

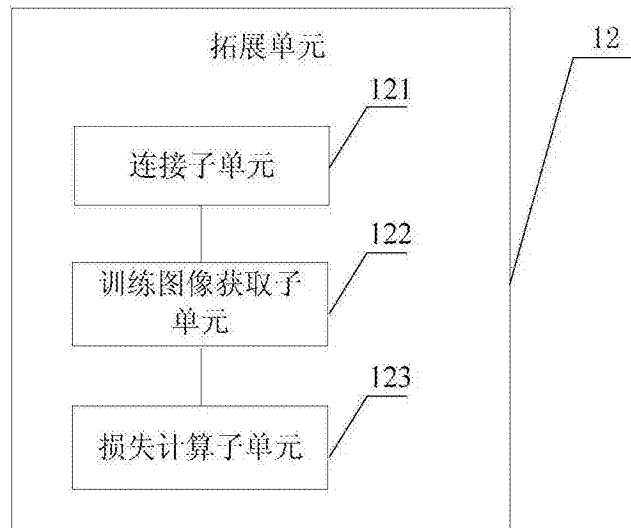


图9

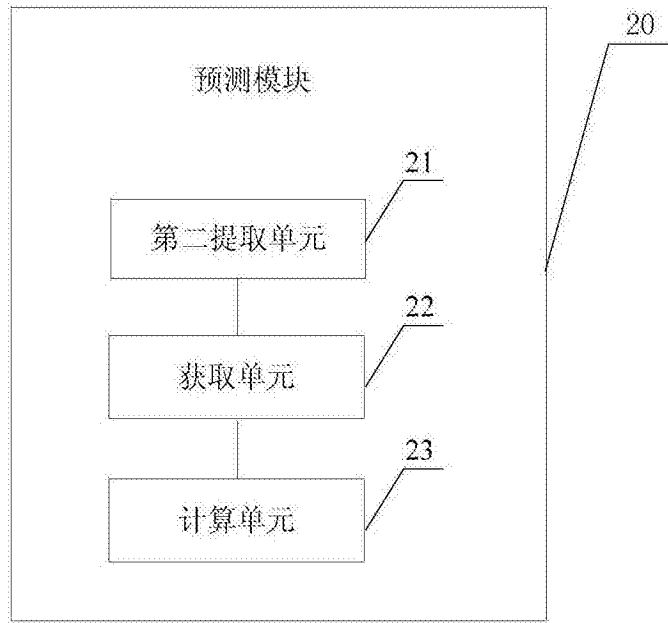


图10