

人脸姿态分析方法、装置、设备、存储介质以及程序

本申请要求在 2017 年 12 月 25 日提交中国专利局、申请号为 201711424986.5、申请名称为“人脸姿态分析方法、装置、设备、存储介质以及程序”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请涉及计算机视觉技术，尤其是涉及一种人脸姿态分析方法、人脸姿态分析装置、神经网络的训练方法、神经网络的训练装置、驾驶状态监控方法、驾驶状态监控装置、电子设备、计算机可读存储介质以及计算机程序。

10

背景技术

人脸图像是包含有人脸的图像。如何从人脸图像中获得更多的信息，是一个值得关注的技术问题。

发明内容

15

本申请实施方式提供一种人脸姿态分析、训练神经网络和驾驶状态监控的技术方案。

根据本申请实施方式其中一方面，提供一种人脸姿态分析方法，所述分析方法包括：获取待处理人脸图像的人脸关键点；将所述待处理人脸图像的人脸关键点输入至神经网络中，以获得所述神经网络输出的所述待处理人脸图像的人脸姿态信息。

20

根据本申请实施方式的其中另一方面，提供一种神经网络的训练方法，所述训练方法包括：从训练数据集中获取训练数据，所述训练数据包括人脸关键点和人脸姿态标注信息；将所述训练数据中的人脸关键点输入待训练的神经网络中，以获得所述待训练的神经网络输出的人脸姿态信息；以所述待训练的神经网络输出的人脸姿态信息与所述训练数据中的人脸姿态标注信息之间的差异为指导信息，对所述待训练的神经网络进行监督学习。

25

根据本申请实施方式的其中再一方面，提供一种驾驶状态监控方法，所述监控方法包括：采用上述人脸姿态分析方法，获得车载摄像头采集的驾驶员图像的人脸姿态信息；根据一段时间内的驾驶员图像的人脸姿态信息，确定驾驶员的驾驶状态；所述驾驶状态包括：分心驾驶状态以及疲劳驾驶状态中的至少一个。

30

根据本申请实施方式的其中再一方面，提供一种人脸姿态分析装置，所述分析装置包括：获取关键点模块，用于获取待处理人脸图像的人脸关键点；获取姿态模块，用于将所述待处理人脸图像的人脸关键点输入至神经网络中，以获得所述神经网络输出的所述待处理人脸图像的人脸姿态信息。

根据本申请实施方式的其中再一方面，提供一种神经网络的训练装置，所述训练装置包括：第一获取模块，用于从训练数据集中获取训练数据；第二获取模块，用于将所述训练数据中的人脸关键点输入待训练的神经网络中，以获得所述待训练的神经网络输出的人脸姿态信息；监督模块，用于以所述待训练的神经网络输出的人脸姿态信息与所述训练数据中的人脸姿态标注信息之间的差异为指导

信息，对所述待训练的神经网络进行监督学习。

根据本申请实施方式的其中再一方面，提供一种驾驶状态监控装置，所述监控装置包括：采用上述人脸姿态分析装置，获得车载摄像头采集的驾驶员图像的人脸姿态信息；确定驾驶状态模块，用于根据一段时间内的驾驶员图像的人脸姿态信息，确定驾驶员的驾驶状态；所述驾驶状态包括：分心驾驶状态以及疲劳驾驶状态中的至少一个。

根据本申请实施方式再一个方面，提供一种电子设备，包括：存储器，用于存储计算机程序；处理器，用于执行所述存储器中存储的计算机程序，且所述计算机程序被执行时，实现本申请任一方法实施方式。

根据本申请实施方式再一个方面，提供一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时，实现本申请任一方法实施方式。

根据本申请实施方式的再一个方面，提供一种计算机程序，包括计算机指令，当所述计算机指令在设备的处理器中运行时，实现本申请任一方法实施方式。

基于本申请提供的人脸姿态分析方法、人脸姿态分析装置、神经网络的训练方法、神经网络的训练装置、电子设备、计算机可读存储介质及计算机程序，本申请通过利用神经网络来提取待处理人脸图像的人脸姿态信息，可以快速准确的获得待处理人脸图像的人脸姿态信息。本申请的人脸姿态分析的技术方案有利于满足实时提取待处理人脸图像的人脸姿态信息的要求，从而有利于使本申请的提取人脸姿态信息的技术方案应用在实时环境中。

应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本申请。下面通过附图和实施例，对本申请的技术方案做进一步的详细描述。

20

附图说明

构成说明书的一部分的附图描述了本申请的实施方式，并且连同描述一起用于解释本申请的原理。参照附图，根据下面的详细描述，可以更加清楚地理解本申请，其中：

图 1 为本申请的人脸姿态分析方法一个实施方式的流程图；

图 2 为本申请的神经网络的训练方法的流程图；

图 3 为本申请的形成训练数据的一个实施方式的流程图；

图 4 为本申请的形成训练数据的另一个实施方式的流程图；

图 5 为本申请的形成训练数据的再一个实施方式的流程图；

图 6 为本申请的人脸姿态分析装置一个实施方式的结构示意图；

图 7 为本申请的神经网络的训练装置一个实施方式的结构示意图；

图 8 为实现本申请实施方式的一示例性设备的框图。

具体实施例

现在将参照附图来详细描述本申请的各种示例性实施例。应注意到：除非另外具体说明，否则在

这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本申请的范围。同时，应当明白，为了便于描述，附图中所示出的各个部分的尺寸并不是按照实际的比例关系绘制的。

以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本申请及其应用或使用的任何限制。对于相关领域普通技术人员已知的技术、方法以及设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为说明书的一部分。

应当注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

本申请实施例可以应用于终端设备、计算机系统及服务器等电子设备，其可与众多其它通用或者专用的计算系统环境或者配置一起操作。适于与终端设备、计算机系统以及服务器等电子设备一起使用的众所周知的终端设备、计算系统、环境和/或配置的例子包括但不限于：个人计算机系统、服务器计算机系统、瘦客户机、厚客户机、手持或膝上设备、基于微处理器的系统、机顶盒、可编程消费电子产品、网络个人电脑、小型计算机系统、大型计算机系统和包括上述任何系统的分布式云计算技术环境，等等。

终端设备、计算机系统以及服务器等电子设备可以在由计算机系统执行的计算机系统可执行指令（诸如程序模块）的一般语境下描述。通常，程序模块可以包括例程、程序、目标程序、组件、逻辑以及数据结构等等，它们执行特定的任务或者实现特定的抽象数据类型。计算机系统/服务器可以在分布式云计算环境中实施，分布式云计算环境中，任务是由通过通信网络链接的远程处理设备执行的。在分布式云计算环境中，程序模块可以位于包括存储设备的本地或远程计算系统存储介质上。

20 示例性实施例

本申请提供的人脸姿态分析方法、神经网络训练的方法和驾驶状态监控方法，可由单片机、FPGA（Field Programmable Gate Array，现场可编程门阵列）、CPU（Central Processing Unit，中央处理单元）、微处理器等数据处理器件，通过调用存储器中存储的相应指令而被执行。

图 1 为本申请人脸姿态分析方法一个实施例的流程图。如图 1 所示，该实施例方法包括：步骤 S100 以及步骤 S110。下面对图 1 中的各步骤进行详细说明。

S100、获取待处理人脸图像的人脸关键点。

在一个可选示例中，本申请中的待处理人脸图像可以为呈现静态的图片或者照片等图像，也可以为呈现动态的视频中的视频帧等；待处理人脸图像可通过摄像头等采集设备预先进行图像采集或边采集的同时边通过本申请技术方案进行实时处理。该待处理人脸图像中的人脸可以为正脸，也可以为存在一定偏转角度的非正脸。本申请中的非正脸的偏转角度可以通过 Pitch（即围绕 x 轴旋转而产生的偏转角）、Yaw（即围绕 y 轴旋转而产生的偏转角）以及 Roll（即围绕 z 轴旋转而产生的偏转角）中的至少一个来表示。本申请不限制待处理人脸图像中的人脸的具体表现形态。

在一个可选示例中，本申请的人脸关键点通常是指用于定位人脸或者人脸局部区域或者一个或多个个人脸器官的多个点，人脸关键点通常包括但不限于人脸轮廓关键点、眼睛关键点、眉毛关键点、嘴

部关键点、鼻子关键点、眼睑线关键点、唇线关键点等。本申请可以利用现有的卷积神经网络来获取待处理人脸图像的人脸关键点，如将待处理人脸图像输入至一卷积神经网络中，根据该卷积神经网络的输出获得待处理人脸图像的人脸关键点。该卷积神经网络的网络结构可以根据提取人脸关键点的实际需求灵活设计，本申请实施例并不限制该卷积神经网络的具体网络结构；例如，本申请的该卷积神经网络可以包括但不限于卷积层、非线性 Relu 层、池化层以及全连接层等，该卷积神经网络所包含的层数越多，则网络越深；再例如，本申请的该卷积神经网络的网络结构可以采用但不限于 AlexNet、深度残差网络（Deep Residual Network, ResNet）或 VGGnet（Visual Geometry Group Network, 视觉几何组网络）等神经网络所采用的网络结构。

在一个可选示例中，本申请获取到的待处理人脸图像的人脸关键点的数量通常不小于 21。在获取到的人脸关键点的数量高于 21 的情况下，本申请最终确定出的待处理人脸图像的人脸姿态信息往往会更准确，例如，利用 106 个人脸关键点所获得的待处理人脸图像的人脸姿态信息的准确性，通常会高于利用 21 个人脸关键点所获得的待处理人脸图像的人脸姿态信息的准确性。本申请不限制获取到的人脸关键点的数量。本申请获取到的待处理人脸图像的人脸关键点通常包括：人脸关键点的编号以及人脸关键点在整个待处理人脸图像或者待处理人脸图像的人脸区域中的二维坐标，即 x 轴坐标和 y 轴坐标。

S110、将待处理人脸图像的人脸关键点输入至神经网络中，以获得神经网络输出的待处理人脸图像的人脸姿态信息。

在一个可选示例中，本申请中的神经网络是利用包含有多条训练数据的训练数据集成功训练出的神经网络。训练数据集中的一条训练数据通常包含有人脸关键点以及人脸姿态标注信息。利用训练数据集训练神经网络的过程可以参见下述针对图 2 的描述。在此不再详细说明。

在一个可选示例中，本申请中的神经网络可以包括：至少两个带有 Relu（Rectified linear unit，修正线性单元）的全连接层以及一个被作为输出层的全连接层；例如，该神经网络包括三层，第一层为带有 Relu 的全连接层，第二层为带有 Relu 的全连接层，第三层为作为输出层的全连接层；再例如，该神经网络包括四层，第一层为带有 Relu 的全连接层，第二层为带有 Relu 的全连接层，第三层为带有 Relu 的全连接层，第四层为作为输出层的全连接层。本申请不限制该神经网络所包含的全连接层的数量。

在一个可选示例中，本申请中的人脸姿态信息是用于描述人脸偏转方向的信息，由于人脸偏转方向可以表征出头部姿态，因此，人脸姿态信息可以称为头部姿态信息。可选的，人脸姿态信息可以包括：围绕 X 轴旋转信息、围绕 Y 轴旋转信息以及围绕 Z 轴旋转信息中的一个或者多个。通常情况下，上述围绕 X 轴旋转信息可以为围绕 X 轴旋转角度（可以称为 Pitch 或者头部俯仰角，表示抬头/低头的角度），上述围绕 Y 轴旋转信息可以为围绕 Y 轴旋转角度（可以称为 Yaw 或者侧脸角度，表示脸部向左/右侧偏转的角度），上述围绕 Z 轴旋转信息可以为围绕 Z 轴旋转角度（可以称为 Roll 或者头部左右翻转角度，表示头顶向左肩/右肩贴近的角度）。在本申请的下述实施例中，以人脸姿态信息包括：Pitch、Yaw 以及 Roll 为例进行描述，然而，这并不表示本申请中的人脸姿态信息必须包括 Pitch、

Yaw 以及 Roll。另外，本申请中的符号“/”的含义表示“或者”。

需要特别说明的是，本申请中的围绕 X 轴旋转信息、围绕 Y 轴旋转信息以及围绕 Z 轴旋转信息也可以通过表征旋转量的其他形式的信息来表示，例如，围绕 X 轴旋转信息可以具体为围绕 X 轴旋转弧度，围绕 Y 轴旋转信息可以具体为围绕 Y 轴旋转弧度，围绕 Z 轴旋转信息可以具体为围绕 Z 轴旋转弧度等，本申请不限制围绕 X 轴旋转信息、围绕 Y 轴旋转信息以及围绕 Z 轴旋转信息具体表现形式。

在一个可选示例中，本申请可以在获得待处理人脸图像的人脸关键点之后，先对待处理人脸图像的人脸关键点进行坐标归一化处理，使人脸关键点的 x 轴坐标和 y 轴坐标的取值范围为[1, -1]，以方便神经网络执行提取人脸姿态信息的操作；然后，再将坐标归一化处理后的脸关键点输入至神经网络中，从而使神经网络基于归一化处理后的脸关键点输出待处理人脸图像的 Pitch、Yaw 以及 Roll。

本申请通过利用神经网络来提取待处理人脸图像的人脸姿态信息，可以快速准确的获得待处理人脸图像的人脸姿态信息，从而有利于避免基于三维人脸模型及梯度下降法优化距离的方式获得待处理人脸图像的人脸姿态信息所具有的计算量大、计算复杂度高以及耗时长等问题，因此，本申请的人脸姿态分析的技术方案有利于满足实时提取待处理人脸图像的人脸姿态信息的要求，从而有利于应用在实时环境中。

本申请提取出的人脸姿态信息有利于对人脸形态的精确勾画，例如，本申请提取出的人脸姿态信息可以用于辅助建立人脸 3D 模型，也可以用于视线分析，还可以用于筛选可见/不可见的人脸关键点等；由此可知，本申请的人脸姿态分析技术可以适用于人脸的图像渲染处理（例如，针对眉毛的特效处理等）、变脸处理（例如，脸部相应器官变形处理）、美颜处理（例如，去皱去斑处理等）、美妆处理（例如，描眉处理等）、人脸识别、人脸状态检测处理（例如，判断眼睛/嘴部的开合程度等）、器官精细定位处理（例如，眼睑线、上/下嘴唇内/外唇线定位等）以及表情检测处理（例如，判断眼睛/嘴部的开合程度等）等应用中。另外，本申请的人脸姿态分析技术还可以适用于确定驾驶状态处理（例如，确定分心驾驶状态、疲劳驾驶状态等）这一应用中。

在一个可选示例中，本申请利用上述人脸姿态分析方法实现驾驶状态监控的过程如下：
首先，获取车载摄像头采集的驾驶员图像（如包含有驾驶员脸部的驾驶员图像），并获取驾驶员图像的人脸关键点，例如，将驾驶员图像输入卷积神经网络，从而获得卷积神经网络输出的驾驶员图像的人脸关键点。

其次，将驾驶员图像的人脸关键点输入至本申请的神经网络中，以获得神经网络输出的驾驶员图像的人脸姿态信息。例如，获得神经网络输出的驾驶员图像的 Pitch、Yaw 以及 Roll。

最后，根据一段时间内的驾驶员图像的人脸姿态信息，确定驾驶员的驾驶状态。例如，在根据一段时间内的驾驶员图像的 Pitch、Yaw 以及 Roll，判断出驾驶员的头部在至少一个方向上的偏转角度和持续偏转时间符合预设条件的情况下，可以确定出驾驶员处于分心驾驶状态或者疲劳驾驶状态。

在一个可选示例中，本申请在确定驾驶状态的过程中，还可以结合驾驶员的手、眼睛以及嘴巴中至少一个，进行判断，从而有利于更准确细致的确定驾驶状态。

一个可选例子，本申请可以根据一段时间内的驾驶员图像的人脸关键点，确定出驾驶员的眼睛开合动作，然后，根据一段时间内的驾驶员图像的人脸姿态信息以及眼睛开合动作，确定出驾驶员的驾驶状态。例如，在根据一段时间内的驾驶员图像判断出驾驶员的头部低垂角度（Pitch）以及低垂持续时间符合预设条件，且眼睛处于闭合或者半闭合状态的持续时间达到一阈值的情况下，确定出驾驶员处于疲劳驾驶状态。

另一个可选例子，本申请可以根据一段时间内的驾驶员图像的人脸关键点，确定出驾驶员的嘴巴开合动作，然后，根据一段时间内的驾驶员图像的人脸姿态信息以及嘴巴开合动作，确定出驾驶员的驾驶状态。例如，在根据一段时间内的驾驶员图像判断出驾驶员的头部左/右偏转角度（Yaw）以及左/右偏转持续时间符合预设条件，且嘴巴开合的次数达到一阈值的情况下，确定出驾驶员处于分心驾驶状态。

再一个可选例子，本申请可以根据一段时间内的驾驶员图像的手部关键点，确定出驾驶员的手在脸部区域的停留状态，然后，根据一段时间内的驾驶员图像的人脸姿态信息以及该停留状态，确定出驾驶员的驾驶状态。例如，在根据一段时间内的驾驶员图像判断出驾驶员的头部向一侧肩膀靠近的侧偏角度（Roll）或者头部低垂角度（Pitch）符合预设条件，且手停留在耳朵或者嘴巴区域附近的持续时间达到一阈值的情况下，确定出驾驶员处于分心驾驶状态。

再一个可选例子，本申请可以根据一段时间内的驾驶员图像的人脸关键点，确定驾驶员的人脸朝向。例如，可以根据可见的人脸关键点和不可见的人脸关键点，确定人脸朝向。然后，本申请根据一段时间内的驾驶员图像的人脸姿态信息以及人脸朝向，确定驾驶员的驾驶状态。例如，在根据一段时间内的驾驶员图像判断出驾驶员的头部左/右偏转角度（Yaw）以及左/右偏转持续时间符合预设条件，且持续人脸朝向也为持续左/右朝向的情况下，确定出驾驶员处于分心驾驶状态。

另外，本申请可以将上述几个可选例子在判断时所使用的判断条件进行组合，例如，在根据一段时间内的驾驶员图像确定出驾驶员的头部向一侧肩膀靠近的侧偏角度（Roll）或者头部低垂角度（Pitch）符合预设条件，手停留在耳朵或者嘴巴区域附近的持续时间达到一阈值，且嘴巴开合的次数达到一阈值的情况下，确定出驾驶员处于使用电话的分心驾驶状态。

在一个可选示例中，本申请响应于驾驶员的驾驶状态满足预定报警条件（如驾驶员处于严重分心驾驶状态或处于严重疲劳驾驶状态），可以执行相应操作，以有利于促进安全驾驶。例如，执行输出报警提示信息的操作。再例如，执行向云端服务器或者预定关联通信端（如驾驶员的移动电话或者驾驶员管理者的移动电话等）发送信息的操作。本申请所发送的信息可以包括：驾驶员的驾驶状态、驾驶员的驾驶状态对应的至少一帧驾驶员图像等。

图2为本申请训练神经网络的一个实施例的流程图。如图2所示，该实施例方法包括：步骤S200、步骤S210以及步骤S220。下面对图2中的各步骤进行详细说明。

S200、从训练数据集中获取训练数据。

在一个可选示例中，本申请中的训练数据集中包括多条用于训练神经网络的训练数据，一条训练数据可以包括：人脸关键点（如21或者36或者106或者160等数量的人脸关键点）以及人脸姿态信

息（如 Pitch、Yaw 以及 Roll）。本申请可以按照随机读取方式或者按照训练数据排列次序顺序读取方式，一次从训练数据集中读取一条或者多条训练数据。训练数据集中的训练数据的生成方式可以如下述针对图 3 的描述，在此不再详细说明。

5 S210、将训练数据中的人脸关键点输入待训练的神经网络中，以获得待训练的神经网络输出的人脸姿态信息。

在一个可选示例中，在训练数据中的人脸关键点为坐标归一化处理后的人脸关键点的情况下，本申请可以直接将人脸关键点输入至待训练的神经网络中，由该神经网络输出围绕 X 轴旋转信息、围绕 Y 轴旋转信息以及围绕 Z 轴旋转信息中的一个或者多个。在训练数据中的人脸关键点不是坐标归一化处理后的人脸关键点的情况下，本申请可以先对人脸关键点进行坐标归一化处理，然后，再将坐标归一化处理后的人脸关键点输入至神经网络中，由该神经网络输出围绕 X 轴旋转信息、围绕 Y 轴旋转信息以及围绕 Z 轴旋转信息中的一个或者多个。通常情况下，该神经网络会针对输入的人脸关键点输出 Pitch、Yaw 以及 Roll。

S220、以神经网络输出的人脸姿态信息和训练数据中的人脸姿态标注信息的差异为指导信息，对待训练的神经网络进行监督学习。

15 在一个可选示例中，本申请可以以减小待训练的神经网络输出的 Pitch、Yaw 和 Roll 与训练数据中的 Pitch、Yaw 以及 Roll 之间的差异为目的，调整待训练的神经网络中的网络参数（如卷积核参数以及权值等）。在达到迭代条件时，本次训练过程结束。本申请的迭代条件可以包括：待训练的神经网络输出的 Pitch、Yaw 和 Roll 与训练数据中的 Pitch、Yaw 以及 Roll 之间的差异满足预定差异要求，且在差异满足该预定差异要求的情况下，本次对待训练的神经网络成功训练完成。本申请中的迭代条件也可以包括对该神经网络进行训练所使用的训练数据的数量达到预定数量要求等。在使用的训练数据的数量达到预定数量要求，然而，差异并未满足该预定差异要求的情况下，本次对待训练的神经网络并未训练成功。成功训练完成的神经网络可以用于从待处理人脸图像中提取人脸姿态信息。

图 3 为本申请形成训练数据集中的训练数据的一个实施例的流程图。如图 3 所示，该实施例方法包括：步骤 S300、步骤 S310 以及步骤 S320。下面对图 3 中的各步骤进行详细说明。

25 S300、获取人脸图像样本的人脸关键点。

在一个可选示例中，本申请可以利用现有的卷积神经网络来获取人脸图像样本的人脸关键点，如将人脸图像样本输入至一卷积神经网络中，根据该卷积神经网络的输出获得人脸图像样本的人脸关键点。另外，本申请获取的人脸图像样本的人脸关键点也可以是针对人脸图像样本人工标注的人脸关键点。

30 S310、根据人脸图像样本的人脸关键点，利用三维人脸模型，确定人脸图像样本的人脸姿态信息。

在一个可选示例中，本申请是基于预先建立的三维人脸模型来确定人脸图像样本的人脸关键点所对应的人脸姿态信息的。该三维人脸模型具有人脸关键点，例如，具有 21 个或者 106 个或者 240 个或者其他数量人脸关键点。本申请可以通过调整三维人脸模型的参数，来提高该三维人脸模型上的人脸关键点的二维投影与上述获取的人脸图像样本的人脸关键点的吻合程度，例如，本申请通过调整三

维人脸模型的参数,可以采用梯度下降法等方法来优化人脸图像样本的人脸关键点和参数调整后的三维人脸模型的人脸关键点投影获得的二维人脸关键点之间的距离之和,从而使该距离之和尽量的小。在上述吻合程度或者上述距离之和符合预定要求的情况下,当前三维人脸模型的人脸姿态信息即为人脸图像样本的人脸姿态信息。

- 5 在一个可选示例中,本申请的调整三维人脸模型的参数通常可以包括调整三维人脸模型的位置以及调整三维人脸模型的姿态(如调整三维人脸模型的 Pitch、Yaw 和 Roll)。进一步的,本申请的调整三维人脸模型的参数还可以包括调整三维人脸模型的表情,即调整三维人脸模型中的与表情相关度较高的人脸关键点,例如,调整位于嘴巴位置处的人脸关键点,再例如,调整位于眼睛位置处的人脸关键点等,以提高三维人脸模型的表情与人脸图像样本的人脸关键点所反映出的表情的相似性。本申请通过调整三维人脸模型的表情,有利于优化人脸图像样本的人脸关键点和参数调整后的三维人脸模型的人脸关键点投影获得的二维人脸关键点之间的距离之和,从而有利于提高人脸姿态标注信息的准确性。

在本申请的调整三维人脸模型的参数包括调整三维人脸模型的位置、调整三维人脸模型的姿态以及调整三维人脸模型的表情,的情况下,本申请的三维人脸模型可以称为精细的三维人脸模型。

- 15 S320、将上述人脸图像样本的人脸姿态信息作为该人脸图像样本的人脸姿态标注信息,与该人脸图像样本的人脸关键点一起,作为一条训练数据添加在训练数据集中。

由于本实施例在形成训练数据的过程中可以借助精细的三维人脸模型获取人脸姿态标注信息,有利于保障人脸姿态标注信息的准确性和可靠性。并且,本实施例仅在神经网络的训练过程中使用三维人脸模型生成训练数据,当神经网络训练完成之后,训练完成的神经网络已经可基于自身在训练过程中学习到的神经元参数来提取人脸姿态信息,一定程度上相当于将精细的三维人脸模型“迁移”到神经网络中,并且处理速度相对于三维人脸模型更快,可更好满足实时分析人脸姿态信息的应用需求。

图4为本申请形成训练数据集中的训练数据的另一实施例的流程图。如图4所示,该实施例方法包括:步骤S400、步骤S410以及步骤S420。下面对图4中的各步骤进行详细说明。

- 25 S400、获取训练数据集中的一条训练数据。获取的该条训练数据通常包含有人脸关键点和人脸姿态标注信息。

S410、对该条训练数据中的人脸关键点进行坐标变换,并根据人脸关键点的坐标变换对该条训练数据中的人脸姿态标注信息进行相应的变换。

- 30 在一个可选示例中,对该条训练数据中的人脸关键点进行坐标变换可以为:设定在人脸为正脸(即人脸正向面对摄像头),人脸关键点围绕Y轴旋转信息和围绕Z轴旋转角信息均表示不存在旋转(如旋转角度为0)的情况下,对该条训练数据中的人脸关键点的X轴坐标进行镜像处理,且在执行镜像处理的情况下,本申请的根据人脸关键点的坐标变换对该条训练数据中的人脸姿态标注信息进行相应的变换可以为:将该条训练数据中的人脸姿态标注信息中的围绕Y轴旋转信息和围绕Z轴旋转信息分别取负,如Yaw和Roll分别取负。

在一个可选示例中,对该条训练数据中的人脸关键点进行坐标变换可以为:对该条训练数据中的

人脸关键点的坐标二维旋转预定角度，且在执行坐标二维旋转预定角度的情况下，本申请的根据人脸关键点的坐标变换对该条训练数据中的人脸姿态标注信息进行相应的变换可以为：根据该预定角度调整该条训练数据中的人脸姿态标注信息中的围绕 Z 轴旋转信息，如 Roll 增加或者减小预定角度。

5 S420、将变换处理后的人脸关键点和人脸姿态标注信息作为一条新的训练数据添加在训练数据集中。

本申请通过对训练数据中的人脸关键点进行坐标变换，并根据人脸关键点的坐标变换对该训练数据中的人脸姿态标注信息进行相应的变换，可以便捷的形成新的训练数据来扩充训练数据集，利用丰富的训练数据来对待训练的神经网络进行训练，有利于避免神经网络的过拟合现象。

10 图 5 为本申请形成训练数据集中的训练数据的再一实施例的流程图。如图 5 所示，该实施例方法包括：步骤 S500、步骤 S510、步骤 S520 以及步骤 S530。下面对图 5 中的各步骤进行详细说明。

S500、将人脸图像样本输入卷积神经网络，以根据卷积神经网络的输出获得该人脸图像样本的人脸关键点。

S510、根据人脸图像样本的人脸关键点，利用三维人脸模型，确定人脸图像样本的 Pitch、Yaw 和 Roll。该步骤所具体执行的操作可以参见上述步骤 S310 的描述。

15 S520、对该人脸图像样本的人脸关键点进行坐标变换，并根据人脸关键点的坐标变换对该人脸图像样本的 Pitch、Yaw 和 Roll 进行相应的变换，获得 Pitch'、Yaw' 和 Roll'。该步骤所具体执行的操作可以参见上述步骤 S410 的描述。

20 S530、将人脸图像样本的人脸关键点以及人脸图像样本的 Pitch、Yaw 和 Roll 作为一条训练数据，将变换处理后的人脸关键点以及 Pitch'、Yaw' 和 Roll' 作为另一条训练数据，分别添加在训练数据集中。本申请通过利用三维人脸模型形成 Pitch、Yaw 和 Roll，有利于保证训练数据的准确性，通过对这样的训练数据进行变换处理，并将变换前后的训练数据均添加在训练数据集中，在有利于保证变换处理后的训练数据的准确性的同时，有利于丰富训练数据集中的训练数据，并有利于提高形成训练数据集的效率。

25 图 6 为本申请的人脸姿态分析装置一个实施例的结构示意图。如图 6 所示，该实施例的装置主要包括：获取关键点模块 600 以及获取姿态模块 610。可选的，该实施例的装置还可以包括：处理模块 620 以及神经网络的训练装置 630。

获取关键点模块 600 主要用于获取待处理人脸图像的人脸关键点。获取关键点模块 600 具体执行的操作，可以参见上述方法实施方式中针对图 1 的 S100 的描述。在此不再重复说明。

30 获取姿态模块 610 主要用于将待处理人脸图像的人脸关键点输入至神经网络中，以获得神经网络输出的待处理人脸图像的人脸姿态信息。获取姿态模块 610 具体执行的操作，可以参见上述方法实施方式中针对图 1 的 S110 的描述。在此不再重复说明。

处理模块 620 主要用于根据待处理人脸图像的人脸姿态信息对待处理人脸图像进行处理。该处理可以包括：人脸识别、变脸处理（例如，脸部部分区域或者整个区域的变形处理）、美颜处理（例如，去皱去斑处理等）、美妆处理（例如，描眉处理等）、人脸的图像渲染处理（例如，针对眉毛的特效

处理等)、人脸状态检测处理(例如,判断眼睛/嘴部的开合程度等)、表情检测处理(例如,判断眼睛/嘴部的开合程度等)、器官精细定位处理(例如,眼睑线、上/下嘴唇内/外唇线定位等)以及确定驾驶状态处理(例如,确定分心驾驶状态、疲劳驾驶状态等)中的至少一个。

5 神经网络的训练装置 630 主要用于利用训练数据集中的训练数据对待训练的神经网络进行训练,其中的训练数据包括人脸关键点和人脸姿态标注信息。神经网络的训练装置 630 具体执行的操作,可以参见上述方法实施方式中针对图 2 的描述。神经网络的训练装置 630 的具体结构可以参见下述实施方式中针对图 7 的描述。在此不再重复说明。

10 在一个可选示例中,本申请可以利用上述人脸姿态分析装置实现驾驶状态监控装置。即驾驶状态监控装置包括:人脸姿态分析装置和确定驾驶状态模块。可选的,驾驶状态监控装置还可以包括:第一确定模块、第二确定模块、第三确定模块、第四确定模块以及响应模块中的至少一个。

人脸姿态分析装置用于获得车载摄像头采集的驾驶员图像的人脸姿态信息。可选的,人脸姿态分析装置获取车载摄像头采集的驾驶员图像(如包含有驾驶员脸部的驾驶员图像),并获取驾驶员图像的人脸关键点,例如,人脸姿态分析装置将驾驶员图像输入卷积神经网络,从而获得卷积神经网络输出的驾驶员图像的人脸关键点。之后,人脸姿态分析装置将其获得的驾驶员图像的人脸关键点输入至神经网络中,以获得神经网络输出的驾驶员图像的人脸姿态信息。例如,人脸姿态分析装置获得神经网络输出的驾驶员图像的 Pitch、Yaw 以及 Roll。

20 确定驾驶状态模块用于根据一段时间内的驾驶员图像的人脸姿态信息,确定驾驶员的驾驶状态。例如,确定驾驶状态模块在根据一段时间内的驾驶员图像的 Pitch、Yaw 以及 Roll,判断出驾驶员的头部在至少一个方向上的偏转角度以及偏转持续时间符合预设条件的情况下,确定出驾驶员处于分心驾驶状态或者疲劳驾驶状态。

25 一个可选例子,第一确定模块可以根据一段时间内的驾驶员图像的人脸关键点,确定出驾驶员的眼睛开合动作,然后,确定驾驶状态模块根据一段时间内的驾驶员图像的人脸姿态信息以及眼睛开合动作,确定出驾驶员的驾驶状态。例如,确定驾驶状态模块在根据一段时间内的驾驶员图像判断出驾驶员的头部低垂角度(Pitch)以及低垂持续时间符合预设条件,且眼睛处于闭合或者半闭合状态的持续时间达到一阈值的情况下,确定出驾驶员处于疲劳驾驶状态。

30 另一个可选例子,第二确定模块可以根据一段时间内的驾驶员图像的人脸关键点,确定出驾驶员的嘴巴开合动作,然后,确定驾驶状态模块根据一段时间内的驾驶员图像的人脸姿态信息以及嘴巴开合动作,确定出驾驶员的驾驶状态。例如,确定驾驶状态模块在根据一段时间内的驾驶员图像判断出驾驶员的头部左/右偏转角度(Yaw)以及左/右偏转持续时间符合预设条件,且嘴巴开合的次数达到一阈值的情况下,确定出驾驶员处于分心驾驶状态。

再一个可选例子,第三确定模块可以根据一段时间内的驾驶员图像的手部关键点,确定出驾驶员的手在脸部区域的停留状态,然后,确定驾驶状态模块根据一段时间内的驾驶员图像的人脸姿态信息和该停留状态,确定出驾驶员的驾驶状态。例如,确定驾驶状态模块在根据一段时间内的驾驶员图像判断出驾驶员的头部向一侧肩膀靠近的侧偏角度(Roll)或头部低垂角度(Pitch)符合预设条件,且

手停留在耳朵或者嘴巴区域附近的持续时间达到一阈值的情况下，确定出驾驶员处于分心驾驶状态。

再一个可选例子，第四确定模块可以根据一段时间内的驾驶员图像的人脸关键点，确定驾驶员的人脸朝向。例如，第四确定模块可以根据可见的人脸关键点和不可见的人脸关键点，确定人脸朝向。然后，确定驾驶状态模块根据一段时间内的驾驶员图像的人脸姿态信息以及人脸朝向，确定驾驶员的
5 驾驶状态。例如，确定驾驶状态模块在根据一段时间内的驾驶员图像判断出驾驶员的头部左/右偏转角度（Yaw）和左/右偏转持续时间符合预设条件，且持续人脸朝向也为持续左/右朝向的情况下，确定出驾驶员处于分心驾驶状态。

另外，确定驾驶状态模块可以将上述不同例子中进行判断时所使用的判断条件进行组合，例如，确定驾驶状态模块在根据一段时间内的驾驶员图像确定出驾驶员的头部向一侧肩膀靠近的侧偏角度
10 （Roll）或者头部低垂角度（Pitch）符合预设条件，手停留在耳朵或者嘴巴区域附近的持续时间达到一阈值，且嘴巴开合的次数达到一阈值的情况下，确定出驾驶员处于使用电话的分心驾驶状态。

响应模块用于响应于驾驶员的驾驶状态满足预定报警条件（如驾驶员处于严重分心驾驶状态或处于严重疲劳驾驶状态），执行相应操作，以有利于促进安全驾驶。例如，响应模块执行输出报警提示信息
15 的操作。再例如，响应模块执行向云端服务器或者预定关联通信端（如驾驶员的移动电话或者驾驶员管理者的移动电话等）发送信息的操作。响应模块发送的信息可以包括：驾驶员的驾驶状态、驾驶员的驾驶状态对应的至少一帧驾驶员图像等。

图7为本申请的神经网络的训练装置一个实施例的结构示意图。如图7所示，该实施例的装置主要包括：第一获取模块700、第二获取模块710以及监督模块720。可选的，该实施例的装置还可以包括：第一标注模块730以及第二标注模块740。

20 第一获取模块700主要用于从训练数据集中获取训练数据。第一获取模块700具体执行的操作，可以参见上述方法实施方式中针对图2的S200的描述。在此不再重复说明。

第二获取模块710主要用于将训练数据中的人脸关键点输入待训练的神经网络中，以获得待训练的神经网络输出的人脸姿态信息。第二获取模块710具体执行的操作，可以参见上述方法实施方式中
针对图2的S210的描述。在此不再重复说明。

25 监督模块720主要用于以待训练的神经网络输出的人脸姿态信息与训练数据中的人脸姿态标注信息之间的差异为指导信息，对待训练的神经网络进行监督学习。监督模块720具体执行的操作，可以参见上述方法实施方式中针对图2的S220的描述。在此不再重复说明。

第一标注模块730主要用于获取人脸图像样本的人脸关键点，并根据人脸图像样本的人脸关键点，利用三维人脸模型，确定人脸图像样本的人脸姿态标注信息；其中，人脸图像样本的人脸关键点
30 和人脸图像样本的人脸姿态标注信息被作为训练数据集中的一条训练数据。第一标注模块730具体执行的操作，可以参见上述方法实施方式中针对图3中的S300-310和图5中的S500-510的描述。在此不再重复说明。

第二标注模块740主要用于获取训练数据集中的一条训练数据，对该条训练数据中的人脸关键点进行坐标变换，并根据人脸关键点的坐标变换对该条训练数据中的人脸姿态标注信息进行相应的变

换；将变换处理后的人脸关键点和人脸姿态标注信息作为一条新的训练数据添加在训练数据集中。第二标注模块 740 具体执行的操作，可以参见上述方法实施方式中针对图 4 中的 S410 和图 5 中的 S520 的描述。在此不再重复说明。

5 示例性设备

图 8 示出了适于实现本申请的示例性设备 800，设备 800 可以是汽车中配置的控制系统/电子系统、移动终端（例如，智能移动电话等）、个人计算机（PC，例如，台式计算机或者笔记型计算机等）、平板电脑以及服务器等。图 8 中，设备 800 包括一个或者多个处理器、通信部等，所述一个或者多个处理器可以为：一个或者多个中央处理单元（CPU）801，和/或，一个或者多个利用神经网络进行人脸姿态分析的图像处理器（GPU）813 等，处理器可以根据存储在只读存储器（ROM）802 中的可执行指令或者从存储部分 808 加载到随机访问存储器（RAM）803 中的可执行指令而执行各种适当的动作和处理。通信部 812 可以包括但不限于网卡，所述网卡可以包括但不限于 IB（Infiniband）网卡。处理器可与只读存储器 802 和/或随机访问存储器 803 中通信以执行可执行指令，通过总线 804 与通信部 812 相连、并经通信部 812 与其他目标设备通信，从而完成本申请中的相应步骤。

上述各指令所执行的操作可以参见上述方法实施例中的相关描述，在此不再详细说明。此外，在 RAM 803 中，还可以存储有装置操作所需的各种程序以及数据。CPU801、ROM802 以及 RAM803 通过总线 804 彼此相连。在有 RAM803 的情况下，ROM802 为可选模块。RAM803 存储可执行指令，或在运行时向 ROM802 中写入可执行指令，可执行指令使中央处理单元 801 执行上述方法所包括的步骤。

输入/输出（I/O）接口 805 也连接至总线 804。通信部 812 可以集成设置，也可以设置为具有多个子模块（例如，多个 IB 网卡），并分别与总线连接。以下部件连接至 I/O 接口 805：包括键盘、鼠标等的输入部分 806；包括诸如阴极射线管（CRT）、液晶显示器（LCD）等以及扬声器等的输出部分 807；包括硬盘等的存储部分 808；以及包括诸如 LAN 卡、调制解调器等网络接口卡的通信部分 809。通信部分 809 经由诸如因特网的网络执行通信处理。驱动器 810 也根据需要连接至 I/O 接口 805。可拆卸介质 811，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 810 上，以便于从其上读出的计算机程序根据需要被安装在存储部分 808 中。

需要特别说明的是，如图 8 所示的架构仅为一种可选实现方式，在具体实践过程中，可根据实际需要对上述图 8 的部件数量和类型进行选择、删减、增加或替换；在不同功能部件设置上，也可采用分离设置或集成设置等实现方式，例如，GPU 和 CPU 可分离设置，再例如，可将 GPU 集成在 CPU 上，通信部可分离设置，也可集成设置在 CPU 或 GPU 上等。这些可替换的实施方式均落入本申请的保护范围。特别地，根据本申请的实施方式，下文参考流程图描述的过程可以被实现为计算机软件程序，例如，本申请实施方式包括一种计算机程序产品，其包含有形地包含在机器可读介质上的计算机程序，计算机程序包含用于执行流程图所示的步骤的程序代码，程序代码可包括对应执行本申请提供的方法中的步骤对应的指令。在这样的实施方式中，该计算机程序可以通过通信部分 809 从网络上被

下载及安装，和/或从可拆卸介质 811 被安装。在该计算机程序被中央处理单元（CPU）801 执行时，执行本申请中记载的实现上述相应步骤的指令。

在一个或多个可选实施方式中，本公开实施例还提供了一种计算机程序产品，用于存储计算机可读指令，所述指令被执行时使得计算机执行上述任意实施例所述的人脸姿态分析方法或者神经网络5 的训练方法或者驾驶状态监控方法。该计算机程序产品可以具体通过硬件、软件或其结合的方式实现。在一个可选例子中，所述计算机程序产品具体体现为计算机存储介质，在另一个可选例子中，所述计算机程序产品具体体现为软件产品，例如软件开发包（Software Development Kit，SDK）等等。

在一个或多个可选实施方式中，本公开实施例还提供了另一种人脸姿态分析方法、神经网络的训练方法和驾驶状态监控方法及其对应的装置和电子设备、计算机存储介质、计算机程序以及计算机程10 序产品，其中，该方法包括：第一装置向第二装置发送人脸姿态分析指示或者训练神经网络指示或者驾驶状态监控指示，该指示使得第二装置执行上述任一可能的实施例所述的人脸姿态分析方法或者训练神经网络方法或者驾驶状态监控方法；第一装置接收第二装置发送的人脸姿态分析结果或者神经网络训练结果或者驾驶状态监控结果。

在一些实施例中，该人脸姿态分析指示或者训练神经网络指示或者驾驶状态监控指示可以具体为15 调用指令，第一装置可以通过调用的方式指示第二装置执行人脸姿态分析操作或者训练神经网络操作或者驾驶状态监控操作，相应地，响应于接收到调用指令，第二装置可以执行上述人脸姿态分析方法或者训练神经网络的方法或者驾驶状态监控方法中的任意实施例中的步骤和/或流程。

应理解，本公开实施例中的“第一”、“第二”等术语仅仅是为了区分，而不应理解成对本公开实施例的限定。还应理解，在本公开中，“多个”可以指两个或两个以上，“至少一个”可以指一个、20 两个或两个以上。

还应理解，对于本公开中提及的任一部件、数据或结构，在没有明确限定或者在前后文给出相反启示的情况下，一般可以理解为一个或多个。

还应理解，本公开对各个实施例的描述着重强调各个实施例之间的不同之处，其相同或相似之处可以相互参考，为了简洁，不再一一赘述。

25 可能以许多方式来实现本申请的方法和装置、电子设备以及计算机可读存储介质。例如，可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本申请的方法和装置、电子设备以及计算机可读存储介质。用于方法的步骤的上述顺序仅是为了进行说明，本申请的方法的步骤不限于以上具体描述的顺序，除非以其它方式特别说明。此外，在一些实施方式中，还可将本申请实施为记录在记录介质中的程序，这些程序包括用于实现根据本申请的方法的机器可读指令。因而，本申请还覆盖存30 储用于执行根据本申请的方法的程序的记录介质。

本申请的描述是为了示例和描述起见而给出的，而并不是无遗漏的或者将本申请限于所公开的形式。很多修改和变化对于本领域的普通技术人员而言，是显然的。选择和描述实施方式是为了更好说明本申请的原理以及实际应用，并且使本领域的普通技术人员能够理解本申请实施例可以从而设计适于特定用途的带有各种修改的各种实施方式。