

驾驶员注意力监测方法和装置及电子设备

本申请要求于 2019 年 3 月 18 日提交中国专利局、申请号为 201910205328.X、发明名称为“驾驶员注意力监测方法和装置及电子设备，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及图像处理技术领域，尤其涉及一种驾驶员注意力监测方法和装置及电子设备。

背景技术

随着道路上的车辆越来越多，如何预防道路交通事故也越来越受关注，其中，人为因素在道路交通事故成因方面占很大的比例，包括由驾驶员注意力不集中、注意力下降等原因引起的分心驾驶。

发明内容

本申请提供一种驾驶员注意力监测技术方案。

第一方面，提供了一种驾驶员注意力监测方法，包括：经车上设置的摄像头针对所述车的驾驶区域采集视频；根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，其中，每帧脸部图像的注视区域属于预先对所述车进行空间区域划分得到的多类定义注视区域之一；根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果。

结合本申请任一实施方式，所述预先对所述车进行空间区域划分得到的多类定义注视区域，包括以下二类或二类以上：左前挡风玻璃区域、右前挡风玻璃区域、仪表盘区域、车内后视镜区域、中控台区域、左后视镜区域、右后视镜区域、遮光板区域、换挡杆区域、方向盘下方区域、副驾驶区域、副驾驶前方的杂物箱区域。

结合本申请任一实施方式，所述根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果，包括：根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述至少一滑动时间窗内各类所述注视区域的注视累计时长；根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，所述注意力监测结果包括是否分心驾驶和/或分心驾驶等级。

结合本申请任一实施方式，所述时间阈值包括：与各类所述定义注视区域分别对应的多个时间阈值，其中，所述多类定义注视区域中至少二个不同类的定义注视区域所对应的时间阈值不同；根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，包括：根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长和相应类别的定义注视区域的时间阈值的比较结果，确定所述驾驶员的注意力监测结果。

结合本申请任一实施方式，所述根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：对所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像进行视线和/或头部姿态检测；根据每帧脸部图像的视线和/或头部姿态的检测结果，确定每帧脸部图像中所述驾驶员的注视区域的类别。

结合本申请任一实施方式，所述根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：将多帧所述脸部图像分别输入神经网络并经所述神经网络分别输出每帧脸部图像中所述驾驶员的注视区域的类别，其中：所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集预先训练完成，或者，所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集以及基于所述人脸图像集中各人脸图像截取的眼部图像预先训练完成；所述注视区域类别标注信息包括所述多类定义注视区域之一。

结合本申请任一实施方式，所述神经网络的训练方法包括：获取所述人脸图像集中包括有注视区域类别标注信息的人脸图像；截取所述人脸图像中的至少一眼的眼部图像，所述至少一眼包括左眼和/或右眼；分别提取所述人脸图像的第一特征和至少一眼的眼部图像

的第二特征；融合所述第一特征和所述第二特征，得到第三特征；根据所述第三特征确定所述人脸图像的注视区域类别检测结果；根据所述注视区域类别检测结果和所述注视区域类别标注信息的差异，调整所述神经网络的网络参数。

结合本申请任一实施方式，所述方法还包括：在所述驾驶员的注意力监测结果为分心驾驶的情况下，对所述驾驶员进行分心驾驶提示，所述分心驾驶提示包括以下至少之一：文字提示、语音提示、气味提示、低电流刺激提示；或者，在所述驾驶员的注意力监测结果为分心驾驶的情况下，根据预先设定的分心驾驶等级与注意监测结果的映射关系、所述驾驶员的注意力监测结果，确定所述驾驶员的分心驾驶等级；根据预先设定的分心驾驶等级与分心驾驶提示的映射关系、所述驾驶员的分心驾驶等级，从所述分心驾驶提示中确定一种提示对所述驾驶员进行分心驾驶提示。

结合本申请任一实施方式，所述预先设定的分心驾驶等级与注意监测结果的映射关系包括：在多个连续滑动时间窗的监测结果均为分心驾驶的情况下，所述分心驾驶等级与滑动时间窗的数量成正相关。

结合本申请任一实施方式，所述经车上设置的摄像头针对所述车的驾驶区域采集视频，包括：经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：根据图像质量评价指标，分别确定采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像中各帧脸部图像的图像质量评分；分别确定所述多个视频时刻对齐的各帧脸部图像中图像质量评分最高的脸部图像；分别确定各图像质量评分最高的脸部图像中所述驾驶员的注视区域的类别。

结合本申请任一实施方式，所述图像质量评价指标包括以下至少之一：图像中是否包括有眼部图像、图像中眼部区域的清晰度、图像中眼部区域的遮挡情况、图像中眼部区域的睁闭眼情况。

结合本申请任一实施方式，所述经车上设置的摄像头针对所述车的驾驶区域采集视频，包括：经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；所述根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：针对采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别检测时刻对齐的各帧脸部图像中所述驾驶员的注视区域类别；将得到的各注视区域类别中多数结果确定为该时刻的脸部图像的注视区域类别。

结合本申请任一实施方式，所述方法还包括：向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果；和/或，对所述驾驶员的注意力监测结果进行统计分析。

结合本申请任一实施方式，在向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果之后，还包括：在接收到所述服务器或所述终端发送的控制指令的情况下，根据所述控制指令控制所述车辆。

第二方面，提供了一种驾驶员注意力监测装置，包括：第一控制单元，用于经车上设置的摄像头针对所述车的驾驶区域采集视频；第一确定单元，用于根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，其中，每帧脸部图像的注视区域属于预先对所述车进行空间区域划分得到的多类定义注视区域之一；第二确定单元，用于根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果。

结合本申请任一实施方式，所述预先对所述车进行空间区域划分得到的多类定义注视区域，包括以下二类或二类以上：左前挡风玻璃区域、右前挡风玻璃区域、仪表盘区域、车内后视镜区域、中控台区域、左后视镜区域、右后视镜区域、遮光板区域、换挡杆区域、方向盘下方区域、副驾驶区域、副驾驶前方的杂物箱区域。

结合本申请任一实施方式，所述第二确定单元包括：第一确定子单元，用于根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述至少一滑动时间窗内各类所述注视区域的注视累计时长；第二确定子单元，用于根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，所述注意力监测结果包括是否分心驾驶和/或分心驾驶等级。

结合本申请任一实施方式，所述时间阈值包括：与各类所述定义注视区域分别对应的多个时间阈值，其中，所述多类定义注视区域中至少二个不同类的定义注视区域所对应的的时间阈值不同；所述第二确定子单元还用于：根据所述至少一滑动时间窗内各类所述注视

区域的注视累计时长和相应类别的定义注视区域的时间阈值的比较结果，确定所述驾驶员的注意力监测结果。

结合本申请任一实施方式，所述第一确定单元包括：第一检测子单元，用于对所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像进行视线和/或头部姿态检测；第三确定子单元，用于根据每帧脸部图像的视线和/或头部姿态的检测结果，确定每帧脸部图像中所述驾驶员的注视区域的类别。

结合本申请任一实施方式，所述第一确定单元还包括：处理子单元，用于将多帧所述脸部图像分别输入神经网络并经所述神经网络分别输出每帧脸部图像中所述驾驶员的注视区域的类别，其中：所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集预先训练完成，或者，所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集以及基于所述人脸图像集中各人脸图像截取的眼部图像预先训练完成；所述注视区域类别标注信息包括所述多类定义注视区域之一。

结合本申请任一实施方式，所述装置还包括所述神经网络的训练单元，所述训练单元包括：获取子单元，用于获取所述人脸图像集中包括有注视区域类别标注信息的人脸图像；图像截取子单元，用于截取所述人脸图像中的至少一眼的眼部图像，所述至少一眼包括左眼和/或右眼；特征提取子单元，用于分别提取所述人脸图像的第一特征和至少一眼的眼部图像的第二特征；特征融合子单元，用于融合所述第一特征和所述第二特征，得到第三特征；第四确定子单元，用于根据所述第三特征确定所述人脸图像的注视区域类别检测结果；调整子单元，用于根据所述注视区域类别检测结果和所述注视区域类别标注信息的差异，调整所述神经网络的网络参数。

结合本申请任一实施方式，所述装置还包括：提示单元，用于在所述驾驶员的注意力监测结果为分心驾驶的情况下，对所述驾驶员进行分心驾驶提示，所述分心驾驶提示包括以下至少之一：文字提示、语音提示、气味提示、低电流刺激提示；第三确定单元，用于在所述驾驶员的注意力监测结果为分心驾驶的情况下，根据预先设定的分心驾驶等级与注意监测结果的映射关系、所述驾驶员的注意力监测结果，确定所述驾驶员的分心驾驶等级；第四确定单元，用于根据预先设定的分心驾驶等级与分心驾驶提示的映射关系、所述驾驶员的分心驾驶等级，从所述分心驾驶提示中确定一种提示对所述驾驶员进行分心驾驶提示。

结合本申请任一实施方式，所述预先设定的分心驾驶等级与注意监测结果的映射关系包括：在多个连续滑动时间窗的监测结果均为分心驾驶的情况下，所述分心驾驶等级与滑动时间窗的数量成正相关。

结合本申请任一实施方式，所述装置还包括：所述第一控制单元，还用于经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；所述第一确定单元，还包括：第五确定子单元，用于根据图像质量评价指标，分别确定采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像中各帧脸部图像的图像质量评分；第六确定子单元，用于分别确定所述多个视频时刻对齐的各帧脸部图像中图像质量评分最高的脸部图像；第七确定子单元，用于分别确定各图像质量评分最高的脸部图像中所述驾驶员的注视区域的类别。

结合本申请任一实施方式，所述图像质量评价指标包括以下至少之一：图像中是否包括有眼部图像、图像中眼部区域的清晰度、图像中眼部区域的遮挡情况、图像中眼部区域的睁闭眼情况。

结合本申请任一实施方式，所述第一控制单元，还用于经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；所述第一确定单元，还包括：第二检测子单元，用于针对采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别检测时刻对齐的各帧脸部图像中所述驾驶员的注视区域类别；第八确定子单元，用于将得到的各注视区域类别中多数结果确定为该时刻的脸部图像的注视区域类别。

结合本申请任一实施方式，所述装置还包括：发送单元，用于向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果；和/或，分析单元，用于对所述驾驶员的注意力监测结果进行统计分析。

结合本申请任一实施方式，所述装置还包括：第二控制单元，用于在向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果之后，且在接收到所述服务器或所述终端发送的控制指令的情况下，根据所述控制指令控制所述车辆。

第三方面，提供了一种电子设备，包括：包括处理器、存储器；所述处理器被配置为支持所述装置执行上述第一方面及其任一种可能的实现方式的方法中相应的功能。存储器

用于与处理器耦合，其保存所述装置必要的程序（指令）和数据。可选的，所述装置还可以包括输入/输出接口，用于支持所述装置与其他装置之间的通信。

第四方面，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中存储有指令，当其在计算机上运行时，使得计算机执行上述第一方面及其任一种可能的实现方式的方法。

第五方面，提供了一种计算机程序产品，所述计算机程序产品包括计算机程序或指令，当所述计算机程序或指令在计算机上运行时，使得计算机执行上述第一方面及其任一种可能的实现方式的方法。

附图说明

此处的附图被并入说明书中并构成本说明书的一部分，这些附图示出了符合本公开的实施例，并与说明书一起用于说明本公开的技术方案。

图1为本申请实施例提供的一种驾驶员注意力监测方法的流程示意图；

图2为本申请实施例提供的一种注视区域划分示意图；

图3为本申请实施例提供的另一种驾驶员注意力监测方法的流程示意图；

图4为本申请实施例提供的一种神经网络的训练方法的流程示意图；

图5为本申请实施例提供的另一种神经网络的训练方法的流程示意图；

图6为本申请实施例提供的另一种驾驶员注意力监测方法的流程示意图；

图7为本申请实施例提供的一种驾驶员注意力监测装置的结构示意图；

图8为本申请实施例提供的一种训练单元的结构示意图；

图9为本申请实施例提供的一种驾驶员注意力监测装置的硬件结构示意图。

具体实施方式

为了使技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请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申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别不同对象，而不是用于描述特定顺序。此外，术语“包括”和“具有”以及它们任何变形，意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、系统、产品或设备没有限定于已列出的步骤或单元，而是可选地还包括没有列出的步骤或单元，或可选地还包括对于这些过程、方法、产品或设备固有的其他步骤或单元。

在本文中提及“实施例”意味着，结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例，也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是，本文所描述的实施例可以与其它实施例相结合。

为了更清楚地说明本申请实施例或背景技术中的技术方案，下面将对本申请实施例或背景技术中所需要使用的附图进行说明。

下面结合本申请实施例中的附图对本申请实施例进行描述。

请参阅图1，图1是本申请实施例提供的一种驾驶员注意力监测方法的流程示意图。

101、经车上设置的摄像头针对所述车的驾驶区域采集视频。

在本申请实施例中，驾驶区域包括车内驾驶室区域。摄像头可安装在车上任意可对驾驶区域进行拍摄的区域，举例来说，摄像头可安装在车内中控台或前挡风玻璃处，也可安装在车后视镜处，还可安装在车的A柱上等等，此外，摄像头的数量可以是一个，也可以是多个，本申请实施例对于摄像头的安装位置以及摄像头的具体数量不做限定。

在一些可能实现的方式中，通过安装在车后视镜处的摄像头对车内驾驶室区域进行视频拍摄，获得驾驶区域的视频。可选地，摄像头可以在接收到某特定指令的情况下，对车的驾驶区域采集视频，例如，将启动车辆（如点火启动、按键启动等等）作为摄像头采集视频的指令，以减少摄像头的能耗；再例如，通过与摄像头连接的终端控制摄像头对驾驶区域采集视频，实现对摄像头的远程控制，可以理解的是，摄像头与终端可通过无线或有线方式进行连接，本申请实施例对于摄像与终端的具体连接方式不做限定。

102、根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，其中，每帧脸部图像的注视区域属于预先对所述车进行空间区域划分得到的多类定义注视区域之一。

在本申请实施例中，驾驶员的脸部图像可以包括驾驶员的整个头部，也可以是包括驾驶员的脸部轮廓以及五官；可以将视频中的任意帧图像作为驾驶员的脸部图像，也可以从视频中任意帧图像中检测出驾驶员的脸部区域图像，并将该脸部区域图像作为驾驶员的脸部图像，上述检测驾驶员脸部区域图像的方式可以是任意人脸检测算法，本申请对此不做具体限定。

在本申请实施例中，可通过将车辆室内空间划分得到的多个不同的区域作为上述多个不同类别的区域，或者可通过将车辆的室外空间划分得到的多个不同的区域作为上述多个不同类别的区域，又或者可通过将车辆室内空间和车辆的室外空间划分得到的多个不同的区域作为上述多个不同类别的注视区域。举例来说，图2为本申请提供的一种注视区域的类别的划分方式，如图2所示，预先对车辆进行空间区域划分得到多类注视区域，包括以下二类或二类以上：左前挡风玻璃区域（1号注视区域）、右前挡风玻璃区域（2号注视区域）、仪表盘区域（3号注视区域）、车内后视镜区域（4号注视区域）、中控台区域（5号注视区域）、左后视镜区域（6号注视区域）、右后视镜区域（7号注视区域）、遮光板区域（8号注视区域）、换挡杆区域（9号注视区域）、方向盘下方区域（10号注视区域）、副驾驶区域（11号注视区域）、副驾驶前方的杂物箱区域（12号注视区域）。采用该方式进行车空间区域的划分，有利于针对性的进行驾驶员的注意力监测；上述方式充分考虑了驾驶员处于驾驶状态时注意力可能落到的各种区域，有利于实现对驾驶员车前向针对性或车前向全空间的注意力监测，由此提高驾驶员注意力监测的准确度和精度。

需要理解的是，由于不同车型的车的空间分布不一样，可根据车型对注视区域的类别进行划分，例如：图2中的驾驶室在车的左侧，正常驾驶时，驾驶员的视线大部分时间在左前挡风玻璃区域，而对于驾驶室在车的右侧的车型，正常驾驶时，驾驶员的视线大部分时间在右前挡风玻璃区域，显然，注视区域的类别的划分应不同于图2中注视区域的类别的划分；此外，还可根据用户的个人喜好对注视区域的类别进行划分，例如：用户觉得中控台的屏幕面积太小，偏好通过屏幕面积更大的终端来控制空调、音响等舒适装置，此时，可根据终端的摆放位置调整注视区域中的中控台区域。还可根据具体情况以其他方式对注视区域的类别的进行划分，本申请对注视区域的类别的划分方式不做限定。

眼睛是驾驶员获取路况信息的主要感觉器官，而驾驶员的视线所在的区域在很大程度上反映了驾驶员的注意力状况，通过对视频包括的多帧位于驾驶区域的驾驶员的脸部图像进行处理，可确定每帧脸部图像中驾驶员的注视区域的类别，进而实现对驾驶员注意力的监测。在一些可能实现的方式中，对驾驶员的脸部图像进行处理，得到脸部图像中驾驶员的视线方向，并根据预先设定的视线方向与注视区域的类别的映射关系，确定脸部图像中驾驶员的注视区域的类别。在另一些可能实现的方式中，对驾驶员的脸部图像进行特征提取处理，根据提取出的特征确定脸部图像中驾驶员的注视区域的类别，一种可选示例中，得到的注视区域的类别为各注视区域对应的预定编号。

103、根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果。

在本申请实施例中，滑动时间窗的大小以及滑动步长可以是预先设置的时长，也可以是脸部图像的数量，在一些可能实现的方式中，滑动时间窗的大小为5秒，滑动步长为0.1秒，设当前时刻下，滑动时间窗的起始时间为10点40分10秒，结束时间为10点40分15秒，则0.1秒后，滑动时间窗的起始时间为10点40分15.1秒，结束时间为10点40分15.1秒，应理解，上述时间均为摄像头采集视频的时间。在另一些可能实现的方式中，对视频中的各帧脸部图像按视频采集视频的时间的先后顺序从小到大进行编号，如：10点40分15秒采集的脸部图像的序号为1，10点40分15.1秒采集的脸部图像的序号为2，依次类推...，设滑动时间窗的大小为10帧脸部图像，滑动步长为1帧脸部图像，设当前时刻下，滑动时间窗内第一帧脸部图像的序号为5，滑动时间窗内最后一帧脸部图像的序号为14，当滑动时间窗前进一个滑动步长后，滑动时间窗内第一帧脸部图像的序号为6，滑动时间窗内最后一帧脸部图像的序号为15。

在本申请一些可选实施例中，注意力监测结果可以包括分心驾驶，或者注意力监测结果可包括疲劳驾驶，又或者注意力监测结果可包括分心驾驶和疲劳驾驶。可选的，注意力监测结果可以包括分心驾驶的等级，或者可包括疲劳驾驶的等级，又或者可包括分心驾驶的等级和疲劳驾驶的等级。由于在驾驶车辆的过程中，驾驶员的视线可能会在不同注视区域内切换，这样，不同时刻采集的脸部图像中驾驶员的注视区域的类别也会相应的改变。以图2为例，正常驾驶时，驾驶员的视线在1号注视区域内的概率较大；因观察路况和车

况的需要, 驾驶员的视线在 2、3、4、6、7 号注视区域内的概率较在 1 号注视区域内的概率要小; 而驾驶员的视线出现在 5、8、9、10、11、12 号注视区域内的概率要比前面两种情况的概率都小; 因此, 通过根据滑动时间窗内各帧脸部图像的注视区域的类别, 确定该滑动时间窗内驾驶员的注视区域的类别分布, 然后根据驾驶员的注视区域的类别确定注意力监测结果。

在一些可能实现的方式中, 以图 2 的注视区域的类别的划分为例, 将 1 号注视区域的第一占比阈值设为 60%; 将 2、3、4、6、7 号注视区域的第二占比阈值设为 40%; 将 5、8、9、10、11、12 号注视区域的第二占比阈值设为 15%; 其中, 当任意一个滑动时间窗内驾驶员的视线在 1 号注视区域内的占比小于或等于 60% 时, 确定注意力监测结果为分心驾驶; 当任意一个滑动时间窗内驾驶员的视线在 2、3、4、6、7 号注视区域内的占比大于或等于 40% 时, 确定注意力监测结果为分心驾驶; 当任意一个滑动时间窗内驾驶员的视线在 5、8、9、10、11、12 号注视区域内的占比大于或等于 15% 时, 确定注意力监测结果为分心驾驶; 若未监测到驾驶员分心驾驶, 确定注意力监测结果为未分心驾驶。如: 一个滑动时间窗内的 10 帧脸部图像中有 4 帧脸部图像的注视区域的类别为 1, 有 3 帧脸部图像的注视区域的类别为 2, 有 2 帧脸部图像的注视区域的类别为 5, 有 1 帧脸部图像的注视区域的类别为 12, 其中, 驾驶员的视线落在 1 号注视区域内的占比为 40%, 驾驶员的视线落在 2、3、4、6、7 号注视区域内的占比为 30%, 驾驶员的视线落在 5、8、9、10、11、12 号注视区域内的占比为 30%, 此时确定驾驶员的注意力监测结果为分心驾驶。在另一些可能实现的方式中, 若在一个滑动时间窗内, 注视区域的类别的分布同时满足以上两种或三种分心驾驶的情况, 注意力监测结果还可包括相应的分心驾驶等级, 可选地, 分心驾驶等级与注视区域的类别的分布满足分心驾驶情况的种类的数量成正相关。

此外, 还可根据连续多个滑动时间窗内所包括的各帧脸部图像的各注视区域的类别分布, 确定驾驶员的注意力监测结果, 在一些可能实现的方式中, 请参见图 2, 正常驾驶时的大部分时间内, 驾驶员的视线在 2 号注视区域内, 由于需要观察路况和车况, 驾驶员的视线也应该出现在 2、3、4、6、7 号注视区域内, 若驾驶员的视线在相当长的一段时间内始终位于 1 号注视区域内, 显然是非正常驾驶状态, 因此, 设置第一阈值, 当驾驶员的视线位于 1 号注视区域内的持续时长达到第一阈值时, 确定驾驶员的注意力监测结果为分心驾驶, 由于滑动时间窗的大小小于第一阈值, 此时可通过连续多个滑动时间窗口内注视区域的类别的分布, 判断驾驶员的视线位于 1 号注视区域内的持续时长是否达到第一阈值。

本申请实施例根据实际需要(例如车型, 例如用户喜好, 例如车型和用户喜好等等), 将车内/外的空间区域划分成不同区域, 得到不同类别的注视区域; 基于摄像头采集的驾驶员的脸部图像, 可确定脸部图像中驾驶员的注视区域的类别; 通过滑动时间窗内的注视区域的类别分布实现对驾驶员注意力的持续监测。该方案通过驾驶员的注视区域的类别监测驾驶员的注意力, 有利于实现对驾驶员车前向针对性或车前向全空间的注意力监测, 由此提高驾驶员注意力监测的精度, 再结合滑动时间窗内的注视区域的类别分布, 进一步提升监测结果的准确度。

请参阅图 3, 图 3 是本申请实施例提供的驾驶员注意力监测方法中步骤 102 的一种可能的实现方式的流程示意图。

301、对所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像进行视线和/或头部姿态检测。

在本申请实施例中, 视线和/或头部姿态检测包括: 视线检测、头部姿态检测、视线检测和头部姿态检测。

通过预先训练好的神经网络对驾驶员的脸部图像进行视线检测和头部姿态检测, 可得到视线信息和/或头部姿态信息, 其中, 视线信息包括视线以及视线的起点位置, 在一种可能实现的方式中, 通过对驾驶员的脸部图像依次进行卷积处理、归一化处理、线性变换, 得到视线信息和/或头部姿态信息。

例如, 可以对驾驶员的脸部图像依次进行驾驶员脸部确认、确定眼部区域、确定虹膜中心, 实现视线检测并确定视线信息。在一些可能实现的方式中, 人在平视或者仰视时眼的轮廓比俯视时要大, 因此首先根据预先测量的眼眶的大小, 将俯视与平视和仰视区分开。然后利用向上看和平视时, 上眼眶到眼睛中心的距离比值的不同, 区分开向上看和平视; 然后再处理向左、中、右看的问题。计算出所有瞳孔点到眼眶左边缘的距离的平方和及右边缘的距离的平方和的比值, 根据该比值确定向左、中、右看时的视线信息。

例如, 可以通过对驾驶员的脸部图像进行处理, 确定驾驶员的头部姿态。在一些可能

实现的方式中，对驾驶员的脸部图像进行面部特征点（如：嘴、鼻子、眼睛）提取，并基于提取出的面部特征点确定脸部图像中面部特征点的位置，再根据面部特征点与头部之间的相对位置，确定脸部图像中的驾驶员的头部姿态。

例如，可同时对视线和头部姿态进行检测，提高检测精度。在一些可能实现的方式中，通过车辆上部署的摄像头采集眼部运动的序列图像，将该序列图像与其正视时的眼部图像进行比较，根据比较的差别得到眼球转过的角度，并基于眼球转过的角度确定视线向量。这里是假设头部不动的情况下得到的检测结果。当头部发生微小转动时，首先建立坐标补偿机制，将正视时的眼部图像进行调整。但是当头部发生较大偏转时，首先观察头部相对于空间某一固定坐标系的变化位置、方向，然后确定视线向量。

可理解，以上为本申请实施例提供的进行视线和/或头部姿态检测的示例，在具体实现中，本领域技术人员还可通过其他方法进行视线和/或头部姿态检测，本申请不做限定。

302、根据每帧脸部图像的视线和/或头部姿态的检测结果，确定每帧脸部图像中所述驾驶员的注视区域的类别。

在本申请实施例中，视线检测结果包括每帧脸部图像中驾驶员的视线向量以及视线向量的起始位置，头部姿态检测结果包括每帧脸部图像中驾驶员的头部姿态，其中，视线向量可以理解为视线的方向，根据视线向量可确定脸部图像中驾驶员的视线相较于驾驶员正视时的视线的偏离角度；头部姿态可以是驾驶员头部在坐标系下的欧拉角，其中，上述坐标系可以为：世界坐标系、相机坐标系、图像坐标系等等。

通过包括有注视区域类别标注信息的视线和/或头部姿态检测结果为训练集对注视区域分类模型进行训练，使训练后的分类模型可根据视线和/或头部姿态的检测结果，确定驾驶员的注视区域的类别，其中，上述注视区域分类模型可以为：决策树分类模型、选择树分类模型、softmax 分类模型等等。在一些可能实现的方式中，视线检测结果和头部姿态检测结果均为特征向量，将视线检测结果和头部姿态检测结果进行融合处理，注视区域分类模型再根据融合后的特征确定驾驶员的注视区域的类别，可选地，上述融合处理可以为特征拼接。在另一些可能实现的方式中，注视区域分类模型可基于视线检测结果或头部姿态检测结果确定驾驶员的注视区域的类别。

不同车型的车内环境以及注视区域的类别的划分方式可能也不相同，在本实施例中，通过与车型对应的训练集对用于对注视区域进行分类的分类器进行训练，可使训练后的分类器适用于不同车型，其中，与车型对应的训练集指包括有该车型注视区域类别标注信息的视线和/或头部姿态检测结果以及对应的新车型的注视区域类别的标注信息，基于训练集对需要在新车型中使用的分类器进行监督训练。分类器可基于神经网络、支持向量机等方式预先构建，本申请对分类器的具体结构不做限定。

例如，在一些可能实现的方式中，A 车型相对驾驶员的前向空间划分为 12 个注视区域，B 车型根据自身的车空间特点，想对驾驶员的前向空间需要相对 A 车型做不同的注视区域划分，如划分成 10 个注视区域。该情形下，基于本实施例构建的驾驶员注意力监控技术方案应用于 A 车型，在需要将该注意力监控技术方案应用于 B 车型之前，可以复用 A 车型中的视线和/或头部姿态检测技术，而只需针对 B 车型的空间特点重新划分注视区域，基于视线和/或头部姿态检测技术以及 B 车型对应的注视区域划分，构建训练集，该训练集包括的人脸图像包括有视线和/或头部姿态检测结果及其对应的 B 车型对应的注视区域的类别标注信息，这样基于构建的训练集对用于 B 车型的注视区域分类的分类器进行监督训练，而无需对用于视线和/或头部姿态检测的模型进行重复训练。训练完成后的分类器以及复用的视线和/或头部姿态检测技术，就构成了本申请实施例提供的驾驶员注意力监测方案。

本实施例将注视区域分类所需的特征信息检测（如视线和/或头部姿态检测）与基于上述特征信息进行注视区域分类，分为两个相对独立的阶段进行，提高了如视线和/或头部姿态等特征信息检测技术在不同车型的复用性，因注视区域划分有变化的新应用场景（如新车型等）仅需对应调整适配新的注视区域划分的分类器或分类方法，减少了因注视区域划分有变化的新应用场景下，驾驶员注意力检测技术方案调整的复杂性和运算量，提高了技术方案的普适性和泛化性，由此更好满足多样化的实际应用需求。

除了将注视区域分类所需的特征信息检测与基于上述特征信息进行注视区域分类，分为两个相对独立的阶段之外，本申请实施例还可基于神经网络实现注视区域类别的端到端的检测，即：向神经网络输入脸部图像，而经神经网络对脸部图像处理输出注视区域类别的检测。其中，神经网络可以基于卷积层、非线性层、全连接层等网络单元按照一定方式堆叠或组成，也可以采用现有的神经网络结构，本申请对此并不限定。确定好待训

训练的神经网络结构之后,所述神经网络可以采用包括有注视区域类别标注信息的人脸图像集进行监督训练,或者,所述神经网络可以采用包括有注视区域类别标注信息的人脸图像集以及基于所述人脸图像集中各人脸图像截取的眼部图像进行监督训练;所述注视区域类别标注信息包括所述多类定义注视区域之一。基于带有上述标注信息的人脸图像集对神经网络进行监督训练,可使得该神经网络能同时学习到进行注视类别区域划分所需的特征提取能力以及注视区域的分类能力,由此实现输入图像输出注视区域类别检测结果的端到端检测。

请参阅图 4,图 4 是本申请实施例提供的用于检测注视区域类别的神经网络的一种可能实现的训练方法的流程示意图。

401、获取包括有所述注视区域类别标注信息的人脸图像集。

在本实施例中,人脸图像集中的每一帧图像都包含注视区域的类别,以图 2 的注视区域的类别的划分为例,每一帧图像包括的标注信息为 1 至 12 中任意一个数字。

402、对所述人脸图像集中的图像进行特征提取处理,获得第四特征。

通过神经网络对人脸图像进行特征提取处理,获得第四特征,在一些可能实现的方式中,对人脸图像依次进行卷积处理、归一化处理、第一线性变换、第二线性变换实现特征提取处理,获得第四特征。

首先,通过神经网络中的多层卷积层对人脸图像进行卷积处理,获得第五特征,其中,每个卷积层提取出的特征内容及语义信息均不一样,具体表现为,通过多层卷积层的卷积处理一步步将图像特征抽象出来,同时也将逐步去除相对次要的特征,因此,越到后面提取出的特征尺寸越小,内容及语义信息就越浓缩。通过多层卷积层逐级对人脸图像进行卷积操作,并提取相应的中间特征,最终得到固定大小的特征数据,这样,可在获得人脸图像主要内容信息(即人脸图像的特征数据)的同时,将图像尺寸缩小,减小系统的计算量,提高运算速度。上述卷积处理的实现过程如下:卷积层对人脸图像做卷积处理,即利用卷积核在人脸图像上滑动,并将人脸图像点上的像素值与对应的卷积核上的数值相乘,然后将所有相乘后的值相加作为卷积核中间像素对应的图像上像素值,最终滑动处理完人脸图像中的所有像素值,并提取出第五特征。需要理解的是,本申请对上述卷积层的数量不做具体限定。

在对人脸图像进行卷积处理时,数据经过每一层网络的处理后,其数据分布都会发生变化,这样会给下一层网络的提取带来困难。因此,在对卷积处理得到的第五特征进行后续处理之前,需要对第五特征进行归一化处理,即将第五特征归一化到均值为 0 且方差为 1 的正态分布。在一些可能实现的方式中,在卷积层后连接归一化处理(batch norm, BN)层,BN 层通过加入可训练的参数对特征进行归一化处理,能加快训练速度,并去除数据的相关性,突出特征之间的分布差异。在一个例子中,BN 层对第五特征的处理过程可参见下文:

假设第五特征为 $\beta = x_{1 \rightarrow m}$, 共 m 个数据,输出是 $y_i = BN(x)$, BN 层将对第五特征进行如下操作:

首先,求出上述第五特征 $\beta = x_{1 \rightarrow m}$ 的平均值,即 $\mu_\beta = \frac{1}{m} \sum_{i=1}^m x_i$;

根据上述平均值 μ_β , 确定上述第五特征的方差,即 $\sigma_\beta^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_\beta)^2$;

根据上述平均值 μ_β 和方差 σ_β^2 , 对上述第五特征进行归一化处理,得到 x_i^- ;

最后,基于缩放变量 γ 和平移变量 δ , 得到归一化的结果,即 $y_i = \gamma x_i^- + \delta$, 其中 γ 和 δ 均为已知。

由于卷积处理以及归一化处理从数据中学习复杂映射的能力较小,无法学习和处理复杂类型的数据,例如图像、视频、音频、语音等等。因此,需要通过对归一化处理后的数据进行线性变换,来解决诸如图像处理、视频处理等复杂问题。在 BN 层后连接线性激活函数,通过激活函数对归一化处理后的数据进行线性变换,可处理复杂的映射,在一些可能实现的方式中,将归一化处理后的数据代入线性整流函数(rectified linear unit, ReLU)函数,实现对归一化处理后的数据的第一线性变换,得到第六特征。

在激活函数层后连接的是全连接（fully connected layers, FC）层，通过全连接层对第六特征进行处理，可将第六特征映射到样本（即注视区域）标记空间。在一些可能实现的方式中，通过全连接层对第六特征进行第二线性变换。全连接层包含输入层（即激活函数层）以及输出层，输出层的任何一个神经元与输入层的每一个神经元都有连接，其中，输出层中的每一个神经元都有相应的权重和偏置，因此，全连接层所有的参数就是每个神经元的权重以及偏置，该权重和偏置的具体大小是通过全连接层进行训练得到的。

将第六特征输入至全连接层时，获取全连接层的权重和偏置（即第二特征数据的权重），再根据权重和偏置对上述第六特征进行加权求和，获得上述第四特征，在一些可能实现的方式中，全连接层的权重和偏置分别为： w_i 和 b_i ，其中 i 为神经元的数量，第六特征为 x ，

则全连接层对第三特征数据进行第二线性变换后得到的第一特征数据为 $\sum_{i=1}^i (w_i x + b_i)$ 。

403、对第一特征数据进行第一非线性变换，获得注视区域类别检测结果。

在全连接层后连接 softmax 层，通过 softmax 层内置的 softmax 函数将输入的不同特征数据映射成 0 至 1 之间的值，且映射后的所有值的和为 1，映射后的值与输入的特征一一对应，这样，就相当于对每个特征数据完成了预测，并以数值的形式给出相应的概率。在一种可能实现的方式中，将第四特征输入至 softmax 层，并将第四特征代入 softmax 函数进行第一非线性变换，获得驾驶员的视线在不同注视区域的概率。

404、根据所述注视区域类别检测结果和所述注视区域类别标注信息的差异，调整所述神经网络的网络参数。

在本实施例中，神经网络包括损失函数，损失函数可以为：交叉熵损失函数、均方差损失函数、平方损失函数等等，本申请对损失函数的具体形式不做限定。

人脸图像集中的每一张图像都有相应的标注信息，即每张人脸图像都会对应一个注视区域类别，将 402 得到的不同注视区域的概率以及标注信息代入损失函数，得到损失函数值。通过调整神经网络的网络参数，使损失函数值小于或等于第二阈值，即可完成神经网络的训练，其中，上述网络参数包括 401 与 402 中各网络层的权重以及偏置。

本实施例根据包括有所述注视区域类别标注信息的人脸图像集对神经网络进行训练，使训练后的神经网络可基于提取出的人脸图像的特征确定注视区域的类别，基于本实施例提供的训练方法只需输入人脸图像集，即可得到训练后的神经网络，训练方式简单，训练时间短。

请参阅图 5，图 5 是本申请实施例提供的上述神经网络的另一种可能实现的训练方法的流程示意图。

501、获取所述人脸图像集中包括有注视区域类别标注信息的人脸图像。

在本实施例中，人脸图像集中的每一张图像都包含注视区域的类别，以图 2 的注视区域的类别的划分为例，每一帧图像包括的标注信息为 1 至 12 中任意一个数字。

通过对不同尺度的特征进行融合，丰富特征信息，可提高注视区域的类别的检测精度，上述丰富特征信息的实现过程可参见 502~505。

502、截取所述人脸图像中的至少一眼的眼部图像，所述至少一眼包括左眼和/或右眼。

上述左眼和/或右眼包括：左眼、右眼、左眼和右眼。

在本实施例中，通过识别人脸图像中的眼部区域图像，并通过截图软件从人脸图像中截取出眼部区域图像，也可以通过画图软件从人脸图像中截取出眼部区域图像等等，本申请对如何识别人脸图像中的眼部区域图像以及如何从人脸图像中截取出眼部区域图像的具体实现方式不做限定。

503、分别提取所述人脸图像的第一特征和至少一眼的眼部图像的第二特征。

在本实施例中，被训练的神经网络包含多个特征提取分支，通过不同的特征提取分支对人脸图像以及眼部图像进行第二特征提取处理，获得人脸图像的第一特征和眼部图像的第二特征，丰富提取出的图像特征尺度，在一些可能实现的方式中，通过不同的特征提取分支分别对人脸图像依次进行卷积处理、归一化处理、第三线性变换、第四线性变换，获得人脸图像特征以及眼部图像特征，其中，视线向量信息包括视线向量以及视线向量的起点位置。应理解，上述眼部图像中可以只包括一只眼睛（左眼或右眼），也可以包括两只眼睛，本申请对此不做限定。

上述卷积处理、归一化处理、第三线性变换、第四线性变换的具体实现过程可参见步骤 402 中的卷积处理、归一化处理、第一线性变换、第二线性变换，此处将不再赘述。

504、融合所述第一特征和所述第二特征，得到第三特征。

由于同一物体（本实施例中指驾驶员）的不同尺度的特征包含的场景信息均不一样，通过将不同尺度的特征进行融合，可得到信息更丰富的特征。

在一些可能实现的方式中，通过对第一特征和第二特征进行融合处理，实现将多个特征中的特征信息融合于一个特征中，有利于提高驾驶员注视区域的类别的检测精度。

505、根据所述第三特征确定所述人脸图像的注视区域类别检测结果。

在本实施例中，注视区域类别检测结果为驾驶员的视线在不同注视区域的概率，取值范围为0至1。在一些可能实现的方式中，将第三特征输入至softmax层，并将第三特征代入softmax函数进行第二非线性变换，获得驾驶员的视线在不同注视区域的概率。

506、根据所述注视区域类别检测结果和所述注视区域类别标注信息的差异，调整所述神经网络的网络参数。

在本实施例中，神经网络包括损失函数，损失函数可以为：交叉熵损失函数、均方差损失函数、平方损失函数等等，本申请对损失函数的具体形式不做限定。

将505得到的不同注视区域的概率以及标注信息代入损失函数，得到损失函数值。通过调整神经网络的网络参数，使损失函数值小于或等于第三阈值，即可完成神经网络的训练，其中，上述网络参数包括503至505中各网络层的权重以及偏置。

通过本实施例提供的训练方式训练得到的神经网络，可对同一帧图像中提取出的不同尺度的特征进行融合，丰富特征信息，进而基于融合后的特征识别驾驶员的注视区域的类别以提高识别精度。

本领域技术人员需理解，本申请提供的两种神经网络的训练方法（401~404以及501~506），可在本地终端（如：电脑或手机或车机端）上实现，也可通过云端实现，本申请对此不做限定。

请参阅图6，图6是本申请实施例提供的驾驶员注意力监测方法中步骤103的一种可能的实现方式的流程图示意图。

601、根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述至少一滑动时间窗内各类所述注视区域的注视累计时长。

在驾驶时，驾驶员的视线在除左前挡风玻璃区域（驾驶室在车的左侧，可参见图2）之外的注视区域内的时长越长，驾驶员分心驾驶的可能性就越大，分心驾驶的等级也越高。因此，根据驾驶员的视线在注视区域的时长，可确定驾驶员的注意力监测结果。由于在驾驶车辆的过程中，驾驶员的视线可能会在不同注视区域内切换，这样，注视区域的类别也会相应的改变。显然，根据驾驶员的视线在注视区域的累计时长确定注意力监测结果，以及根据驾驶员的视线在注视区域的连续时长确定注意力监测结果，均是不合理的，因此，通过滑动时间窗对驾驶员的注意力进行监测，实现对驾驶员注意力的持续监测。首先，根据滑动时间窗内每一帧脸部图像的注视区域的类别以及每一帧脸部图像的时长，确定该滑动时间窗内每个注视区域的累计时长。在一些可能实现的方式中，以图2的注视区域的类别的划分为例，一个滑动时间窗内的10帧脸部图像中有4帧脸部图像的注视区域的类别为1，3帧脸部图像的注视区域的类别为2，2帧脸部图像的注视区域的类别为5，1帧脸部图像的注视区域的类别为12，一帧脸部图像的时长为0.4秒，则在该滑动时间窗内，1号注视区域的累计时长为1.6秒，2号注视区域的累计时长为1.2秒，5号注视区域的0.8秒，12号注视区域的累计时长为0.4秒。

602、根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，所述注意力监测结果包括是否分心驾驶和/或分心驾驶等级。

本申请实施中，分心驾驶和/或分心驾驶等级包括：分心驾驶、分心驾驶等级、分心驾驶和分心驾驶等级。

如上所述，由于驾驶的需要，在一定时间内驾驶员的注视区域的类别可能会有多个，显然，不同的注视区域对应分心驾驶的概率均不一样，以图2为例，当驾驶员的注视区域为1时，驾驶员分心驾驶的概率较小，当驾驶员的注视区域为10时，驾驶员分心驾驶的概率较大。因此，通过为不同类别的注视区域设置不同的时间阈值，以体现驾驶员的视线在不同类别的注视区域时，驾驶员分心驾驶的概率不同。再根据至少一滑动时间窗内各类注视区域的注视累计时长和相应类别的定义注视区域的时间阈值的比较结果，确定驾驶员的注意力监测结果，这样，每个滑动时间窗对应一个注意力监测结果。

可选地，在一个滑动时间窗内驾驶员的视线在任意一个注视区域内的累计时长达到该

注视区域的时间阈值时，确定驾驶员的注意力检测结果为分心驾驶。在一些可能实现的方式中，以图 2 为例，设滑动时间窗的时长为 5 秒，驾驶员在需要观察右前方的路况时，视线会在注视区域 2 内；在驾驶过程中，驾驶员需要通过观察仪表盘显示的数据了解车辆的实时状况时，视线会在注视区域 3 内；正常驾驶时，驾驶员的视线不应该出现在注视区域 10 内，因此，可将注视区域 2、3、10 的时间阈值分别设置为：2.5 秒、1.5 秒、0.7 秒；检测到在一个滑动时间窗内，驾驶员的注视区域的类别为 2、3、10 的累计时长分别为 1.8 秒、1 秒、1 秒，则驾驶员的注意力检测结果为分心驾驶。需要理解的是，滑动时间窗的大小以及注视区域的时间阈值的大小，可根据实际使用情况进行调整，本申请对此不做具体限定。

可选地，注意力监测结果还包括分心驾驶等级，即当连续多个滑动时间窗的注意力监测结果均为分心驾驶时，对应的分心驾驶等级也会相应地提高，如：任意一个滑动时间窗的注意力监测结果为分心驾驶对应地分心驾驶等级为 1 级，连续 2 个滑动时间窗的注意力监测结果为分心驾驶对应地分心驾驶等级为 2 级。

可选地，可在车辆室内的不同地方部署多个摄像头，或者可在车辆的室外的不同的地方部署多个摄像头，又或者可在车辆室内和车辆室外的不同地方部署多个摄像头。通过上述多个摄像头可获得同一时刻下的多张人脸图像，经过处理每一帧人脸图像都将有一个注视区域的类别，此时将综合每一帧图像的注视区域的类别确定驾驶员的注视区域的类别，为此，本实施例提供一种“少数服从多数”的投票方法，以确定注视区域的类别，由此提高注视区域类别检测的可靠性，进而提高驾驶员注意力检测的准确性。该方法包括以下步骤：

经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；

针对采集到的多个视频各自包括的多帧位于驾驶区域的驾驶员的脸部图像，分别检测时刻对齐的各帧脸部图像中驾驶员的注视区域类别；

将得到的各注视区域类别中多数结果确定为该时刻的脸部图像的注视区域类别。

在本实施例中，多个视频时刻对齐的各帧脸部图像指：多个摄像头采集的视频中，同一时刻的各帧脸部图像。在一些可能实现的方式中，车上部署有 3 个摄像头，分别为 1 号摄像头、2 号摄像头、3 号摄像头，且可通过这 3 个摄像头分别从不同角度对驾驶区域进行视频采集，可将这 3 个摄像头分别安装在车的不同位置，以从不同角度对驾驶区域进行视频采集，等等。例如，在同一时刻下，1 号摄像头采集到的人脸图像对应的注视区域的类别为右前挡风玻璃区域、2 号摄像头采集到的人脸图像对应的注视区域的类别为车内后视镜区域、3 号摄像头采集到的人脸图像对应的注视区域的类别为右前挡风玻璃区域，由于 3 个结果中有 2 个结果为右前挡风玻璃区域，只有一个结果为车内后视镜区域，所以最终输出的驾驶员的注视区域为右前挡风玻璃区域，注视区域的类别为 2。

可选地，现实环境的光线复杂，汽车内的光线复杂程度更甚，而光照强度会直接影响摄像头的拍摄质量，低质量的图像或视频会丢失部分有用的信息。此外，不同的拍摄角度对拍摄的图像的质量也有影响，导致视频或图像中特征物不明显或被遮挡等问题，如：因驾驶员的眼镜镜片反光导致摄像头无法将驾驶员的眼睛拍清楚，或者，因驾驶员的头部姿态导致未能拍摄到眼睛部分的图像，进而影响后续基于图像实现的检测处理。为此，本实施例还提供一种基于多角度拍摄的图像中挑选质量好的图像作为用于进行驾驶员注视区域类别检测的图像，由于作为检测基础的图像的质量有所保证，因此提高注视区域类别检测的准确性，为不同光线环境、人脸大角度或遮挡等场景提高了解决方案，提高了驾驶员注意力监测的准确性。该方法包括以下步骤：

经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；

根据图像质量评价指标，分别确定采集到的多个视频各自包括的多帧位于驾驶区域的驾驶员的脸部图像中各帧脸部图像的图像质量评分；

分别确定多个视频时刻对齐的各帧脸部图像中图像质量评分最高的脸部图像；

分别确定各图像质量评分最高的脸部图像中驾驶员的注视区域的类别。

在本实施例中，图像质量评价指标包括以下至少之一：图像中是否包括有眼部图像、图像中眼部区域的清晰度、图像中眼部区域的遮挡情况、图像中眼部区域的睁闭眼情况；多个视频时刻对齐的各帧脸部图像指：多个摄像头采集的视频中，同一时刻的各帧脸部图像。通过上述图像质量评价指标确定出的图像可更准确地进行图像中驾驶员注视区域的检测。

在一些可能实现的方式中，在同一时刻，部署在车辆不同地方的摄像头分别从不同角度获取包含驾驶员脸部的图像，根据上述图像质量评价指标对所有图像的质量进行评分，

如：若图像中包含有眼部图像获得 5 分，再根据图像中眼部区域的清晰度从 1~5 分中获得相应的分数，最后将两项得到相加，得到图像质量评分，将同一时刻下不同角度的摄像头采集的多帧图像中图像质量评分最高的图像作为该时刻用于确定注视区域类别的待处理图像，并确定待处理图像中驾驶员的注视区域的类别。需要理解的是，图像中眼部区域的清晰度的判断可由任意图像清晰度算法实现，如：灰度方差函数、灰度方差乘积函数、能量梯度函数，对此本申请不做具体限定。

本实施例根据滑动时间窗内各类注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员是否分心驾驶；根据滑动时间窗的数量确定分心驾驶的等级；通过部署在车上不同区域的摄像头从多个角度对驾驶区域进行视频采集，提高采集到的人脸图像的图像质量，并通过图像质量评价指标确定图像质量最高的人脸图像，基于图像质量最高的人脸图像确定注意力监测结果可提高监测精度；对于车上部署有多个摄像头的情况，还根据“少数服从多数”原则，从同一时刻下多个摄像头对应的多个注意力监测结果中确定注意力监测结果，同样可提高检测精度。

在确定驾驶员分心驾驶的情况下，可及时对驾驶员进行提示，使驾驶员专心驾驶，以下实施例本申请提供的分心驾驶提示的一种可能实现的方式。

在驾驶员的注意力监测结果为分心驾驶的情况下，可对驾驶员进行相应地分心驾驶提示，使驾驶员专心驾驶。分心驾驶提示包括以下至少之一：文字提示、语音提示、气味提示、低电流刺激提示。

在一些可能实现的方式中，当检测到驾驶员的注意力监测结果为分心驾驶时，通过抬头数字显示仪（head up display，HUD）显示器弹出对话框，对驾驶员发出提示和警告；还可通过车载终端内置的语音数据进行提示和警告，如：“请驾驶员专心驾驶”；还可通过释放具有的醒脑提神功效的气体，如：通过车载喷头喷出花露水喷雾，花露水的气味清香怡人，在对驾驶员进行提示和警告的同时，还能起到醒脑提神的效果；还可通座椅释放出低电流刺激驾驶员，以达到提示和警告的效果。

本实施例提供了几种分心驾驶提示方式，以实现在驾驶员分心驾驶的情况下，对驾驶员进行有效的提示和警告。

以下实施例本申请提供的分心驾驶提示的另一种可能实现的方式。

如上所述，当连续多个滑动时间窗的注意力监测结果均为分心驾驶时，对应的分心驾驶等级也会相应地提高，在所述驾驶员的注意力监测结果为分心驾驶的情况下，根据预先设定的分心驾驶等级与注意监测结果的映射关系、所述驾驶员的注意力监测结果，确定所述驾驶员的分心驾驶等级；根据预先设定的分心驾驶等级与分心驾驶提示的映射关系、所述驾驶员的分心驾驶等级，从所述分心驾驶提示中确定一种提示对所述驾驶员进行分心驾驶提示，其中，预先设定的分心驾驶等级与注意监测结果的映射关系包括：在多个连续滑动时间窗的监测结果均为分心驾驶的情况下，分心驾驶等级与滑动时间窗的数量成正相关。

在一些可能实现的方式中，滑动时间窗的数量、分心驾驶等级以及提示方式的映射关系可参见表 1。

滑动时间窗的数量（个）	分心驾驶等级	提示方式
1	1	气味提示
2 或 3	2	文字提示
4 或 5	3	语音提示
6 至 8	4	低电流刺激提示
大于或等于 9	5	语音提示+低电流刺激提示

表 1

当任意一个滑动时间窗的注意力监测结果为分心驾驶时，确定驾驶员的分心驾驶等级为 1，此时，通过气味提示的方式对驾驶员进行提示和警告，如：释放具有的醒脑提神功效的气体，如：通过车载喷头喷出花露水喷雾；当连续 2 个或 3 个滑动时间窗的注意力监测结果为分心驾驶时，确定驾驶员的分心驾驶等级为 2，此时，通过文字提示的方式对驾驶员进行提示和警告，如：通过 HUD 显示器弹出对话框，对驾驶员发出提示和警告；当连续 4 个或 5 个滑动时间窗的注意力监测结果为分心驾驶时，确定驾驶员的分心驾驶等级为 3，此时，通过语音提示的方式对驾驶员进行提示和警告，如：车载终端发出“请驾驶员专心驾驶”的提示语句；当连续 6 个至 8 个滑动时间窗的注意力监测结果为分心驾驶时，确定驾驶员的分心驾驶等级为 4，此时，通过低电流刺激提示的方式对驾驶员进行提示和警

告,如:从驾驶员座椅释放出低电流刺激驾驶员;当连续9个或以上滑动时间窗的注意力监测结果为分心驾驶时,确定驾驶员的分心驾驶等级为5,此时,对驾驶员同时进行语音提示和低电流刺激提示,提示驾驶员专心驾驶。

本实施例根据滑动时间窗的数量、分心驾驶等级以及提示方式的映射关系确定驾驶员分心驾驶的等级,并给出不同程度的提示,实现以合理的方式及时对驾驶员进行提示,使驾驶员专心驾驶,预防因驾驶员分心驾驶导致的交通事故。

在确定驾驶员的注意力监测结果后,可对驾驶员的注意力监测结果进行分析,如:根据驾驶员的注意力监测结果,确定驾驶员的驾驶习惯,并给出产生分心驾驶的原因。还可将注意力监测结果发送给服务器或终端,相关人员可通过服务器或终端实现对车辆的远程控制,或基于注意力监测结果获知驾驶员的驾驶状态,并基于驾驶员的驾驶状态进行相应的处理。以下实施例为本申请提供的基于注意力监测结果的一些可能实现的方式。

车辆可与服务器或终端建立通信连接,其中,上述通信连接可以为蜂窝网络连接、近距离无线通讯(near field communication, NFC)连接、蓝牙连接等等,本申请对通信连接的方式不做限定。在确定驾驶员注意力监测结果的情况下,向与车辆通信连接的服务器或终端发送驾驶员的注意力监测结果,使服务器侧的相关人员以及终端侧的使用者实时掌控驾驶员的注意力监控结果。

在一些可能实现的方式中,物流公司的相关工作人员可通过服务器实时获知每个驾驶员的注意力监测结果,也可对服务器储存的驾驶员的注意力监测结果进行统计,并根据统计结果对驾驶员进行管理。在一些可能实现的方式中,物流公司C规定将驾驶员在物流运输过程中的注意力监测结果作为驾驶员的考核依据之一,例如:在任意一次物流运输过程中,分心驾驶的累计时间占物流运输总时间的比例大于或等于5%,考核评分减1分;分心驾驶的累计时间占物流运输总时间的比例大于或等于7%,考核评分减2分;分心驾驶的累计时间占物流运输总时间的比例大于或等于10%,考核评分减3分;分心驾驶的累计时间占物流运输总时间的比例小于或等于3%,考核评分加1分;分心驾驶的累计时间占物流运输总时间的比例小于或等于2%,考核评分加2分;分心驾驶的累计时间占物流运输总时间的比例小于或等于1%,考核评分加3分。又例如:每出现一次1级分心驾驶,考核评分减0.1分;每出现一次2级分心驾驶,考核评分减0.2分;每出现一次3级分心驾驶,考核评分减0.3分;每出现一次4级分心驾驶,考核评分减0.4分;每出现一次5级分心驾驶,考核评分减0.5分。

进一步地,基于对驾驶员的管理可以对车队进行管理,在另一些可能实现的方式中,物流公司C可根据驾驶员的考核评分对驾驶员进行评级,考核评分越高,相应地,级别也越高。显然,驾驶员的级别越高,驾驶员的驾驶习惯相对也越好,其中,驾驶习惯可以为:不分心驾驶,不疲劳驾驶等等,而对于优先级较高的运输任务,物流公司C可优先指定级别的驾驶员运输,这样,既能确保运输任务的顺利完成,也可让驾驶员信服公司的安排。

车辆通过NFC或蓝牙与车内其他人员(除驾驶员之外的任意一个人)的移动终端(如:手机、平板电脑、笔记本电脑、可穿戴设备等)连接,并实时将驾驶员的注意力监测结果发送至该移动终端,这样,车内其他人员可在驾驶员分心驾驶时对驾驶员进行提醒。在一些可能实现的方式中,丈夫是驾驶员,妻子坐在副驾驶座椅上用平板电脑看电影,妻子通过平板电脑上弹出的消息获知丈夫正分心驾驶,且分心驾驶等级已达到3级,此时,妻子可放下手中的平板电脑对丈夫进行口头提醒,如:“眼睛看哪里呢,专心开车!”,这样可对丈夫起到提醒和警告的作用,使丈夫专心驾驶。通过终端显示驾驶员的注意力监测结果的方式不局限于上述“弹出”,还可以是语音提示,动态效果显示等等,本申请对此不做限定。需要理解的是,在这种实现方式中,车内其他人员可结合注意力监测结果、路况、车况等因素人为判断是否需要提醒,或者需要对驾驶员进行什么程度的提醒,显然,在绝大多数情况下,人的判断能力要优于机器的判断能力,因此,通过车内其他人员给予驾驶员的提示的效果较表1中的提示方式要更好。

通过蜂窝网络将驾驶员的注意力监测结果发送给与车辆通信连接的终端,其中,终端可以为可移动终端,也可以为不可移动终端,终端使用者可以为驾驶员的家人,也可以是驾驶员信任的人,本申请对此不做限定。终端使用者可根据驾驶员的注意力监测结果,采取相应的措施,预防交通事故的发生。在一些可能实现的方式中,在家里的父亲通过手机获知身为驾驶员的儿子正分心驾驶,分心驾驶等级已达到5级,且注意力监测结果为分心驾驶的滑动时间窗的数量还在一直增加,显然,驾驶员的驾驶状态非常不正常,极易发生交通事故,此时,父亲可打电话给正坐在副驾驶座椅上看电影的儿媳,让她对儿子进行

提醒，或采取其他措施，以降低安全隐患。

可选地，也可通过终端向车辆发送控制指令，如：切换驾驶模式，或调整报警模式，或既切换驾驶模式又调整报警模式等等，在接收到服务器或终端发送的控制指令的情况下，根据控制指令控制车辆，在一些可能实现的方式中，通过车辆的远程控制终端向车辆发送控制指令，将车辆的驾驶模式由非自动驾驶模式切换为自动驾驶模式，以使车辆在自动驾驶模式下自动驾驶，减少因驾驶员的不安全驾驶带来的安全隐患。在另一些可能实现的方式中，通过车辆的远程控制终端向车辆发送控制指令，以调整车辆的报警模式（如调大车辆上的报警器的音量等等），以增强报警效果，由此降低安全隐患。在又一些可能实现的方式中，通过车辆的远程控制终端向车辆发送控制指令，既将车辆的驾驶模式由非自动驾驶模式切换为自动驾驶模式又调整车辆的报警模式。

车载终端还可对驾驶员的注意力检测结果进行统计分析，获得分析结果，如：发生分心驾驶的时间、分心驾驶的次数、分心驾驶的累计时间、每一次分心驾驶的等级、驾驶员的驾驶习惯信息，其中，驾驶习惯信息包括分心驾驶时的注视区域的类别分布、产生分心驾驶的原因，在一些可能实现的方式中，车载终端对驾驶员的注意力监测结果进行统计，得到分心驾驶时的注视区域的类别分布，如：以图2为例，在刚刚过去的一周内，分心驾驶时，50%的注视区域的类别是12号区域，30%的注视区域的类别是7号区域，10%的注视区域的类别是2号区域，10%的注视区域的类别是其他区域。进一步地，可根据注视区域的类别分布给出驾驶员分心驾驶的原因，如：驾驶时与副驾驶座椅上的乘客交谈。将注视区域的类别分布以及产生分心驾驶的原因以统计报告的形式呈现给驾驶员，使驾驶员及时获知自身驾驶习惯，并做出相应地调整。可选地，还可将发生分心驾驶的时间、分心驾驶的次数、分心驾驶的累计时间、每一次分心驾驶的等级的统计结果以报表的形式呈现给驾驶员。应用本实施例，可将驾驶员的注意力监测结果发送至服务器并进行储存，相关人员可通过服务器储存的注意力监测结果实现对驾驶员的管理；通过将驾驶员的注意力监测结果发送至车内其他终端，使车内其他人员及时获知驾驶员的驾驶状态，并对驾驶员进行相应的提醒，预防交通事故的发生；通过将驾驶员的注意力监测结果发送至远程终端，可使其他人员根据注意力监测结果对车辆进行相应地控制，降低安全隐患；通过对驾驶员的注意力监测结果进行分析，驾驶员可根据分析结果更清楚的了解自身驾驶状态，并及时矫正自身的不良驾驶习惯，预防交通事故的发生。

本领域技术人员可以理解，在具体实施方式的上述方法中，各步骤的撰写顺序并不意味着严格的执行顺序而对实施过程构成任何限定，各步骤的具体执行顺序应当以其功能和可能的内在逻辑确定。

请参阅图7，图7为本申请实施例提供的一种识别分心驾驶的装置的结构示意图，该装置1包括：第一控制单元11、第一确定单元12、第二确定单元13、提示单元14、第三确定单元15、第四确定单元16、训练单元17、发送单元18、分析单元19以及第二控制单元20，其中：

第一控制单元11，用于经车上设置的摄像头针对所述车的驾驶区域采集视频；以及在车上多个区域分别部署从不同角度的摄像头，并经多个摄像头分别采集驾驶区域的视频流；以及用于经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；

第一确定单元12，用于根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，其中，每帧脸部图像的注视区域属于预先对所述车进行空间区域划分得到的多类定义注视区域之一；以及在车上多个区域分别部署从不同角度的摄像头，并经多个摄像头分别采集驾驶区域的视频流，对采集到的多个视频流分别检测同一时刻的脸部图像中的注视区域类别；

第二确定单元13，用于根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果；

提示单元14，用于在所述驾驶员的注意力监测结果为分心驾驶的情况下，对所述驾驶员进行分心驾驶提示，所述分心驾驶提示包括以下至少之一：文字提示、语音提示、气味提示、低电流刺激提示；

第三确定单元15，用于在所述驾驶员的注意力监测结果为分心驾驶的情况下，根据预先设定的分心驾驶等级与注意监测结果的映射关系、所述驾驶员的注意力监测结果，确定所述驾驶员的分心驾驶等级；

第四确定单元16，用于根据预先设定的分心驾驶等级与分心驾驶提示的映射关系、所述驾驶员的分心驾驶等级，从所述分心驾驶提示中确定一种提示对所述驾驶员进行分心驾

驶提示；

训练单元 17，用于训练所述神经网络；

发送单元 18，用于向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果；

分析单元 19，用于对所述驾驶员的注意力监测结果进行统计分析；

第二控制单元 20，用于在向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果之后，且在接收到所述服务器或所述终端发送的控制指令的情况下，根据所述控制指令控制所述车辆。

在一种可能实现的方式中，所述预先对所述车进行空间区域划分得到的多类定义注视区域，包括以下二类或二类以上：左前挡风玻璃区域、右前挡风玻璃区域、仪表盘区域、车内后视镜区域、中控台区域、左后视镜区域、右后视镜区域、遮光板区域、换挡杆区域、方向盘下方区域、副驾驶区域、副驾驶前方的杂物箱区域。

进一步地，所述第二确定单元 13 包括：第一确定子单元 131，用于根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述至少一滑动时间窗内各类所述注视区域的注视累计时长；第二确定子单元 132，用于根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，所述注意力监测结果包括是否分心驾驶和/或分心驾驶等级。

进一步地，所述时间阈值包括：与各类所述定义注视区域分别对应的多个时间阈值，其中，所述多类定义注视区域中至少二个不同类的定义注视区域所对应的的时间阈值不同；所述第二确定子单元 132 还用于：根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长和相应类别的定义注视区域的时间阈值的比较结果，确定所述驾驶员的注意力监测结果。

进一步地，所述第一确定单元 12 包括：第一检测子单元 121，用于对所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像进行视线和/或头部姿态检测；第三确定子单元 122，用于根据每帧脸部图像的视线和/或头部姿态的检测结果，确定每帧脸部图像中所述驾驶员的注视区域的类别。

进一步地，所述第一确定单元 12 还包括：处理子单元 123，用于将多帧所述脸部图像分别输入神经网络并经所述神经网络分别输出每帧脸部图像中所述驾驶员的注视区域的类别，其中：所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集预先训练完成，或者，所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集以及基于所述人脸图像集中各人脸图像截取的眼部图像预先训练完成；所述注视区域类别标注信息包括所述多类定义注视区域之一。

进一步地，所述预先设定的分心驾驶等级与注意监测结果的映射关系包括：在多个连续滑动时间窗的监测结果均为分心驾驶的情况下，所述分心驾驶等级与滑动时间窗的数量成正相关。

进一步地，所述第一确定单元 12，还包括：第五确定子单元 124，用于根据图像质量评价指标，分别确定采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像中各帧脸部图像的图像质量评分；第六确定子单元 125，用于分别确定所述多个视频时刻对齐的各帧脸部图像中图像质量评分最高的脸部图像；第七确定子单元 126，用于分别确定各图像质量评分最高的脸部图像中所述驾驶员的注视区域的类别。

进一步地，所述图像质量评价指标包括以下至少之一：图像中是否包括有眼部图像、图像中眼部区域的清晰度、图像中眼部区域的遮挡情况、图像中眼部区域的睁闭眼情况。

进一步地，所述第一确定单元 12，还包括：第二检测子单元 127，用于针对采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别检测时刻对齐的各帧脸部图像中所述驾驶员的注视区域类别；第八确定子单元 128，用于将得到的各注视区域类别中多数结果确定为该时刻的脸部图像的注视区域类别。

请参阅图 8，图 8 为本申请实施例提供的一种训练单元 17 的结构示意图，该单元 17 包括：获取子单元 171，用于获取所述人脸图像集中包括有注视区域类别标注信息的人脸图像；图像截取子单元 172，用于截取所述人脸图像中的至少一眼的眼部图像，所述至少一眼包括左眼和/或右眼；特征提取子单元 173，用于分别提取所述人脸图像的第一特征和至少一眼的眼部图像的第二特征；特征融合子单元 174，用于融合所述第一特征和所述第二特征，得到第三特征；第四确定子单元 175，用于根据所述第三特征确定所述人脸图像的注视区域类别检测结果；调整子单元 176，用于根据所述注视区域类别检测结果和所述

注视区域类别标注信息的差异，调整所述神经网络的网络参数。

在一些实施例中，本公开实施例提供的装置具有的功能或包含的模块可以用于执行上文方法实施例描述的方法，其具体实现可以参照上文方法实施例的描述，为了简洁，这里不再赘述。

图9为本申请实施例提供的一种驾驶员注意力监测装置的硬件结构示意图。该监测装置3包括处理器31，还可以包括输入装置32、输出装置33和存储器34。该输入装置32、输出装置33、存储器34和处理器31之间通过总线相互连接。

存储器包括但不限于是随机存储记忆体（random access memory，RAM）、只读存储器（read-only memory，ROM）、可擦除可编程只读存储器（erasable programmable read only memory，EPROM）、或便携式只读存储器（compact disc read-only memory，CD-ROM），该存储器用于相关指令及数据。

输入装置用于输入数据和/或信号，以及输出装置用于输出数据和/或信号。输出装置和输入装置可以是独立的器件，也可以是一个整体的器件。

处理器可以包括是一个或多个处理器，例如包括一个或多个中央处理器（central processing unit，CPU），在处理器是一个CPU的情况下，该CPU可以是单核CPU，也可以是多核CPU。

存储器用于存储网络设备的程序代码和数据。

处理器用于调用该存储器中的程序代码和数据，执行上述方法实施例中的步骤。具体可参见方法实施例中的描述，在此不再赘述。

可以理解的是，图9仅仅示出了一种驾驶员注意力监测装置的简化设计。在实际应用中，驾驶员注意力监测装置还可以分别包含必要的其他元件，包括但不限于任意数量的输入/输出装置、处理器、控制器、存储器等，而所有可以实现本申请实施例的驾驶员注意力监测装置都在本申请的保护范围之内。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。所属领域的技术人员还可以清楚地了解到，本申请各个实施例描述各有侧重，为描述的方便和简洁，相同或类似的部分在不同实施例中可能没有赘述，因此，在某一实施例未描述或未详细描述的部分可以参见其他实施例的记载。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者通过所述计算机可读存储介质进行传输。所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（digital subscriber line，DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质

或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，(例如，软盘、硬盘、磁带)、光介质(例如，数字通用光盘(digital versatile disc, DVD))、或者半导体介质(例如固态硬盘(solid state disk, SSD))等。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，该流程可以由计算机程序来指令相关的硬件完成，该程序可存储于计算机可读取存储介质中，该程序在执行时，可包括如上述各方法实施例的流程。而前述的存储介质包括：只读存储器

(read-only memory, ROM)或随机存取存储器(random access memory, RAM)、磁碟或者光盘等各种可存储程序代码的介质。

权利要求

- 1、一种驾驶员注意力监测方法，其特征在于，包括：
经车上设置的摄像头针对所述车的驾驶区域采集视频；
根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，其中，每帧脸部图像的注视区域属于预先对所述车进行空间区域划分得到的多类定义注视区域之一；
根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果。
- 2、根据权利要求1所述的方法，其特征在于，所述预先对所述车进行空间区域划分得到的多类定义注视区域，包括以下二类或二类以上：左前挡风玻璃区域、右前挡风玻璃区域、仪表盘区域、车内后视镜区域、中控台区域、左后视镜区域、右后视镜区域、遮光板区域、换挡杆区域、方向盘下方区域、副驾驶区域、副驾驶前方的杂物箱区域。
- 3、根据权利要求1或2所述的方法，其特征在于，所述根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果，包括：
根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述至少一滑动时间窗内各类所述注视区域的注视累计时长；
根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，所述注意力监测结果包括是否分心驾驶和/或分心驾驶等级。
- 4、根据权利要求3所述的方法，其特征在于，所述时间阈值包括：与各类所述定义注视区域分别对应的多个时间阈值，其中，所述多类定义注视区域中至少二个不同类的定义注视区域所对应的时间阈值不同；
根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，包括：根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长和相应类别的定义注视区域的时间阈值的比较结果，确定所述驾驶员的注意力监测结果。
- 5、根据权利要求1至4任意一项所述的方法，其特征在于，所述根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：
对所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像进行视线和/或头部姿态检测；
根据每帧脸部图像的视线和/或头部姿态的检测结果，确定每帧脸部图像中所述驾驶员的注视区域的类别。
- 6、根据权利要求1至4任意一项所述的方法，其特征在于，所述根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：
将多帧所述脸部图像分别输入神经网络并经所述神经网络分别输出每帧脸部图像中所述驾驶员的注视区域的类别，其中：所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集预先训练完成，或者，所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集以及基于所述人脸图像集中各人脸图像截取的眼部图像预先训练完成；所述注视区域类别标注信息包括所述多类定义注视区域之一。
- 7、根据权利要求6所述的方法，其特征在于，所述神经网络的训练方法包括：
获取所述人脸图像集中包括有注视区域类别标注信息的人脸图像；
截取所述人脸图像中的至少一眼的眼部图像，所述至少一眼包括左眼和/或右眼；
分别提取所述人脸图像的第一特征和至少一眼的眼部图像的第二特征；
融合所述第一特征和所述第二特征，得到第三特征；

根据所述第三特征确定所述人脸图像的注视区域类别检测结果；
根据所述注视区域类别检测结果和所述注视区域类别标注信息的差异，调整所述神经网络的网络参数。

8、根据权利要求1至7任意一项所述的方法，其特征在于，所述方法还包括：

在所述驾驶员的注意力监测结果为分心驾驶的情况下，对所述驾驶员进行分心驾驶提示，所述分心驾驶提示包括以下至少之一：文字提示、语音提示、气味提示、低电流刺激提示；或者，

在所述驾驶员的注意力监测结果为分心驾驶的情况下，根据预先设定的分心驾驶等级与注意监测结果的映射关系、所述驾驶员的注意力监测结果，确定所述驾驶员的分心驾驶等级；根据预先设定的分心驾驶等级与分心驾驶提示的映射关系、所述驾驶员的分心驾驶等级，从所述分心驾驶提示中确定一种提示对所述驾驶员进行分心驾驶提示。

9、根据权利要求1至8任意一项所述的方法，其特征在于，所述预先设定的分心驾驶等级与注意监测结果的映射关系包括：在多个连续滑动时间窗的监测结果均为分心驾驶的情况下，所述分心驾驶等级与滑动时间窗的数量成正相关。

10、根据权利要求1至9任意一项所述的方法，其特征在于，所述经车上设置的摄像头针对所述车的驾驶区域采集视频，包括：经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；

根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：根据图像质量评价指标，分别确定采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像中各帧脸部图像的图像质量评分；分别确定所述多个视频时刻对齐的各帧脸部图像中图像质量评分最高的脸部图像；分别确定各图像质量评分最高的脸部图像中所述驾驶员的注视区域的类别。

11、根据权利要求10所述的方法，其特征在于，所述图像质量评价指标包括以下至少之一：图像中是否包括有眼部图像、图像中眼部区域的清晰度、图像中眼部区域的遮挡情况、图像中眼部区域的睁闭眼情况。

12、根据权利要求1至9任意一项所述的方法，其特征在于，所述经车上设置的摄像头针对所述车的驾驶区域采集视频，包括：经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；

所述根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，包括：针对采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别检测时刻对齐的各帧脸部图像中所述驾驶员的注视区域类别；将得到的各注视区域类别中多数结果确定为该时刻的脸部图像的注视区域类别。

13、根据权利要求1至12任意一项所述的方法，其特征在于，所述方法还包括：
向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果；和/或，
对所述驾驶员的注意力监测结果进行统计分析。

14、根据权利要求13所述的方法，其特征在于，在向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果之后，还包括：

在接收到所述服务器或所述终端发送的控制指令的情况下，根据所述控制指令控制所述车辆。

15、一种驾驶员注意力监测装置，其特征在于，包括：

第一控制单元，用于经车上设置的摄像头针对所述车的驾驶区域采集视频；

第一确定单元，用于根据所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别确定每帧脸部图像中所述驾驶员的注视区域的类别，其中，每帧脸部图像的注视区域属于预先对所述车进行空间区域划分得到的多类定义注视区域之一；

第二确定单元，用于根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述驾驶员的注意力监测结果。

16、根据权利要求 15 所述的装置，其特征在于，所述预先对所述车进行空间区域划分得到的多类定义注视区域，包括以下二类或二类以上：左前挡风玻璃区域、右前挡风玻璃区域、仪表盘区域、车内后视镜区域、中控台区域、左后视镜区域、右后视镜区域、遮光板区域、换挡杆区域、方向盘下方区域、副驾驶区域、副驾驶前方的杂物箱区域。

17、根据权利要求 15 或 16 所述的装置，其特征在于，所述第二确定单元包括：

第一确定子单元，用于根据所述视频中至少一滑动时间窗内所包括的各帧脸部图像的各所述注视区域的类别分布，确定所述至少一滑动时间窗内各类所述注视区域的注视累计时长；

第二确定子单元，用于根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长与预定的时间阈值的比较结果，确定所述驾驶员的注意力监测结果，所述注意力监测结果包括是否分心驾驶和/或分心驾驶等级。

18、根据权利要求 17 所述的装置，其特征在于，所述时间阈值包括：与各类所述定义注视区域分别对应的多个时间阈值，其中，所述多类定义注视区域中至少二个不同类的定义注视区域所对应的时间阈值不同；

所述第二确定子单元还用于：根据所述至少一滑动时间窗内各类所述注视区域的注视累计时长和相应类别的定义注视区域的时间阈值的比较结果，确定所述驾驶员的注意力监测结果。

19、根据权利要求 15 至 18 任意一项所述的装置，其特征在于，所述第一确定单元包括：

第一检测子单元，用于对所述视频包括的多帧位于所述驾驶区域的驾驶员的脸部图像进行视线和/或头部姿态检测；

第三确定子单元，用于根据每帧脸部图像的视线和/或头部姿态的检测结果，确定每帧脸部图像中所述驾驶员的注视区域的类别。

20、根据权利要求 15 至 18 任意一项所述的装置，其特征在于，所述第一确定单元还包括：

处理子单元，用于将多帧所述脸部图像分别输入神经网络并经所述神经网络分别输出每帧脸部图像中所述驾驶员的注视区域的类别，其中：所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集预先训练完成，或者，所述神经网络预先采用包括有注视区域类别标注信息的人脸图像集以及基于所述人脸图像集中各人脸图像截取的眼部图像预先训练完成；所述注视区域类别标注信息包括所述多类定义注视区域之一。

21、根据权利要求 20 所述的装置，其特征在于，所述装置还包括所述神经网络的训练单元，所述训练单元包括：

获取子单元，用于获取所述人脸图像集中包括有注视区域类别标注信息的人脸图像；

图像截取子单元，用于截取所述人脸图像中的至少一眼的眼部图像，所述至少一眼包括左眼和/或右眼；

特征提取子单元，用于分别提取所述人脸图像的第一特征和至少一眼的眼部图像的第二特征；

特征融合子单元，用于融合所述第一特征和所述第二特征，得到第三特征；

第四确定子单元，用于根据所述第三特征确定所述人脸图像的注视区域类别检测结果；

调整子单元，用于根据所述注视区域类别检测结果和所述注视区域类别标注信息的差异，调整所述神经网络的网络参数。

22、根据权利要求 15 至 21 任意一项所述的装置，其特征在于，所述装置还包括：

提示单元，用于在所述驾驶员的注意力监测结果为分心驾驶的情况下，对所述驾驶员进行分心驾驶提示，所述分心驾驶提示包括以下至少之一：文字提示、语音提示、气味提

示、低电流刺激提示；

第三确定单元，用于在所述驾驶员的注意力监测结果为分心驾驶的情况下，根据预先设定的分心驾驶等级与注意监测结果的映射关系、所述驾驶员的注意力监测结果，确定所述驾驶员的分心驾驶等级；

第四确定单元，用于根据预先设定的分心驾驶等级与分心驾驶提示的映射关系、所述驾驶员的分心驾驶等级，从所述分心驾驶提示中确定一种提示对所述驾驶员进行分心驾驶提示。

23、根据权利要求 15 至 22 任意一项所述的装置，其特征在于，所述预先设定的分心驾驶等级与注意监测结果的映射关系包括：在多个连续滑动时间窗的监测结果均为分心驾驶的情况下，所述分心驾驶等级与滑动时间窗的数量成正相关。

24、根据权利要求 15 至 23 任意一项所述的装置，其特征在于，所述第一控制单元，还用于经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；

所述第一确定单元，还包括：第五确定子单元，用于根据图像质量评价指标，分别确定采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像中各帧脸部图像的图像质量评分；

第六确定子单元，用于分别确定所述多个视频时刻对齐的各帧脸部图像中图像质量评分最高的脸部图像；

第七确定子单元，用于分别确定各图像质量评分最高的脸部图像中所述驾驶员的注视区域的类别。

25、根据权利要求 24 所述的装置，其特征在于，所述图像质量评价指标包括以下至少之一：图像中是否包括有眼部图像、图像中眼部区域的清晰度、图像中眼部区域的遮挡情况、图像中眼部区域的睁闭眼情况。

26、根据权利要求 15 至 23 任意一项所述的装置，其特征在于，所述第一控制单元，还用于经在车上多个区域分别部署的多个摄像头从不同角度分别采集驾驶区域的视频；

所述第一确定单元，还包括：第二检测子单元，用于针对采集到的多个视频各自包括的多帧位于所述驾驶区域的驾驶员的脸部图像，分别检测时刻对齐的各帧脸部图像中所述驾驶员的注视区域类别；

第八确定子单元，用于将得到的各注视区域类别中多数结果确定为该时刻的脸部图像的注视区域类别。

27、根据权利要求 15 至 26 任意一项所述的装置，其特征在于，所述装置还包括：发送单元，用于向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果；和/或，

分析单元，用于对所述驾驶员的注意力监测结果进行统计分析。

28、根据权利要求 27 所述的装置，其特征在于，所述装置还包括：

第二控制单元，用于在向与所述车辆通信连接的服务器或终端发送所述驾驶员的注意力监测结果之后，且在接收到所述服务器或所述终端发送的控制指令的情况下，根据所述控制指令控制所述车辆。

29、一种电子设备，其特征在于，包括存储器和处理器，所述存储器上存储有计算机可执行指令，所述处理器运行所述存储器上的计算机可执行指令时实现权利要求 1 至 14 任一项所述的方法。

30、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中存储有计算机程序，该计算机程序被处理器执行时，实现权利要求 1 至 14 任一项所述的方法。

31、一种计算机程序产品，其特征在于，所述计算机程序产品包括计算机程序或指令，当所述计算机程序或指令在计算机上运行时，实现权利要求 1 至 14 任一项所述的方法。