

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021年2月4日 (04.02.2021)



(10) 国际公布号
WO 2021/016873 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2019/098407
- (22) 国际申请日: 2019年7月30日 (30.07.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 珠海全志科技股份有限公司 (ALLWINNER TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省珠海市高新区唐家湾镇科技二路9号, Guangdong 519085 (CN)。
- (72) 发明人: 李晓会 (LI, Xiaohui); 中国广东省珠海市高新区唐家湾镇科技二路9号, Guangdong 519085 (CN)。 彭刚 (PENG, Gang); 中国广东省珠海市高新区唐家湾镇科技二路9号, Guangdong 519085 (CN)。 南楠 (NAN, Nan); 中国广东省珠海市高新区唐家湾镇科技二路9号, Guangdong 519085 (CN)。 叶丽萍 (YE, Liping);
- (74) 代理人: 珠海智专专利商标代理有限公司 (INNOPAT INTELLECTUAL PROPERTY CO., LTD.); 中国广东省珠海市南屏坪岚路南屏企业大厦第六层, Guangdong 519060 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,

(54) **Title:** CASCADED NEURAL NETWORK-BASED ATTENTION DETECTION METHOD, COMPUTER DEVICE, AND COMPUTER-READABLE STORAGE MEDIUM

(54) **发明名称:** 基于级联神经网络的注意力检测方法、计算机装置及计算机可读存储介质

$$P_i = \frac{e^i}{\sum_j e^j}, i = 1, 2, \dots, n$$

图 2

(57) **Abstract:** Provided in the present invention are a cascaded neural network-based attention detection method, a computer device, and a computer-readable storage medium. The method comprises: acquiring video data, recognizing multiple images, extracting facial area of the multiple images; moreover, applying a first convolutional neural network in recognizing the facial area, determining whether a first scenario of inattention is presented; if a determination is made that no first scenario of inattention is presented, then applying a second convolutional neural network in recognizing the facial area, and determining whether a second scenario of inattention is presented, where the computational complexity of the first convolutional neural network is less than the computational complexity of the second convolutional neural network. Also provided in the present invention are the computer device and the computer-readable storage medium for implementing the method.

(57) **摘要:** 本发明提供一种基于级联神经网络的注意力检测方法、计算机装置及计算机可读存储介质, 该方法包括获取视频数据, 并对多帧图像进行识别, 提取多帧图像的人脸区域; 并且, 应用第一卷积神经网络对人脸区域进行识别, 判断是否出现注意力不集中的第一情形; 如确认没有出现注意力不集中的第一情形, 则应用第二卷积神经网络对人脸区域进行识别, 判断是否出现注意力不集中的第二情形; 其中, 第一卷积神经网络的计算复杂度小于第二卷积神经网络的计算复杂度。本发明还提供实现上述方法的计算机装置及计算机可读存储介质。



WO 2021/016873 A1

NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

根据细则4.17的声明:

- 发明人资格(细则4.17(iv))

本国际公布:

- 包括国际检索报告(条约第21条(3))。

基于级联神经网络的注意力检测方法、计算机装置及计算机可读存储介质

技术领域

[0001] 本发明涉及图像识别的处理领域，具体地，是基于级联神经网络的注意力检测方法以及实现这种方法的计算机装置、计算机可读存储介质。

背景技术

[0002] 随着智能技术的发展，通过图像识别的方式来对人的注意力检测已经成为前沿的新兴技术。人的注意力检测一直以来是机器学习领域研究的重点和热点之一，其主要应用于安防、辅助驾驶等方面。由于真实环境存在大量不确定的因素，如白天、黑夜等不同的光照条件的影响，又例如人的头部姿态和表情具有多样性，还存在人种、性别和年龄等差异以及人穿戴眼镜等因素，因此，真实环境下检测人的注意力状态是相当具有挑战性的。

[0003] 有鉴于此，如何提高人的注意力检测性能成为人工智能研究领域的热点，人们也为此提出了多种算法。例如，远程眼球跟踪是一种检测人的注意力的经典算法，在室外环境下该方法需要依赖于近红外照明设备来产生明亮的瞳孔效果，进而捕捉眼球信息，但是，这种方法所使用的近红外照明设备受到震动和颠簸的影响，很容易被损坏，需长期进行维护，成本比较高。

[0004] 因此，一些研究人员提出依靠头部姿态和眼部信息实现人的注意力的检测。一种方案是将头部姿态、眼部特征以及汽车的几何特征进行融合，进而对人眼关注的区域进行分类，实现人的注意力的检测，该方法取得了很好的效果。而另一种头部姿态和眼部特征相结合的方法是对人眼关注的区域进行分类。但上述两种方法有两个主要问题，第一个是需进行人脸检测、人脸标定、眼球检测以及特征提取等一系列复杂操作，如果该算法中的某一个子模块性能不佳，势必会影响整体效果；第二个是在进行特征提取时，采用传统机器学习方法与传统特征提取的算法泛化能力较差，例如在摄像头拍摄角度、外界光照条件以及目标的位置发生变化时，该方法性能急剧下降。

[0005] 因此，一些研究人员提出一种基于卷积神经网络的注意力估计方法，该方法能够自动从数据样本中学习头部姿态特征和眼部特征信息，无需手动设计特征提取算法，鲁棒性好，但采用的卷积神经网络的模型体积大，计算复杂度高，不适用于嵌入式设备，导致这种方法的使用受到很大的限制。

发明概述

技术问题

[0006] 本发明的主要目的是提供一种计算复杂度小且计算性能好的基于级联神经网络的注意力检测方法。

[0007] 本发明的另一目的是提供一种实现上述基于级联神经网络的注意力检测方法的计算机装置。

[0008] 本发明的再一目的是提供一种实现上述基于级联神经网络的注意力检测方法的计算机可读存储介质。

技术解决手段

[0010] 为实现本发明的主要目的，本发明提供的基于级联神经网络的注意力检测方法包括获取视频数据，并对多帧图像进行识别，提取多帧图像的人脸区域；并且，应用第一卷积神经网络对人脸区域进行识别，判断是否出现注意力不集中的第一情形；如确认没有出现注意力不集中的第一情形，则应用第二卷积神经网络对人脸区域进行识别，判断是否出现注意力不集中的第二情形；其中，第一卷积神经网络的计算复杂度小于第二卷积神经网络的计算复杂度。

[0011] 一个优选的方案是，应用第二卷积神经网络对人脸区域进行识别包括：从人脸区域划截取出多种感兴趣区域，根据二种以上的感兴趣区域的识别结果判断是否出现注意力不集中的第二情形。

[0012] 进一步的方案是，多种感兴趣区域包括人脸框区域以及人脸补充区域；根据二种以上的感兴趣区域的识别结果判断是否出现注意力不集中的第二情形包括：根据人脸框区域以及人脸补充区域的图像识别结果判断是否出现注意力不集中的第二情形。

[0013] 更进一步的方案是，多种感兴趣区域包括人脸框区域以及眼部区域；根据二种以上的感兴趣区域的识别结果判断是否出现注意力不集中的第二情形包括：根

据人脸框区域以及眼部区域的图像识别结果判断是否出现注意力不集中的第二情形。

[0014] 更进一步的方案是，判断是否出现注意力不集中的第一情形包括：应用第一卷积神经网络对人脸区域进行识别，判断头部向预设方向的转动角度是否大于预设角度，如是，确认出现注意力不集中的第一情形。

[0015] 更进一步的方案是，第二卷积神经网络包括依次级联的第一卷积层、逐深度卷积层、多个瓶颈残差层、第二卷积层、线性全局逐深度卷积层、线性卷积层、全连接层以及分类层。

[0016] 更进一步的方案是，瓶颈残差层包括卷积单元以及接收卷积单元输出的逐深度卷积单元，还设置有残差单元，残差单元在卷积单元的步长为1时，实现瓶颈残差层的残差运算。

[0017] 更进一步的方案是，获取视频数据后对多帧图像进行识别包括：从视频数据的每连续预设帧数的图像中选取一帧图像进行识别。

[0018] 为实现上是的另一目的，本发明提供的计算机装置包括处理器以及存储器，存储器存储有计算机程序，计算机程序被处理器执行时实现上述基于级联神经网络的注意力检测方法的各个步骤。

[0019] 为实现上是的再一目的，本发明提供计算机可读存储介质上存储有计算机程序，计算机程序被处理器执行时实现上述基于级联神经网络的注意力检测方法的各个步骤。

问题的解决方案

发明的有益效果

有益效果

[0020] 应用本发明的方案，在提取多帧图像的人脸区域以后，首先通过第一级卷积神经网络进行识别，并判断是否出现注意力不集中的第一情形，只有确认没有出现注意力不集中的第一情形时，才使用第二卷积神经网络进行识别，判断是否出现注意力不集中的第二情形。这样，可以避免针对所有情况都使用计算程度比较复杂的卷积神经网络进行计算，从而简化注意力检测的整体复杂程度。

[0021] 此外，本发明的方案应用第二卷积神经网络对是否出现注意力不集中的第二情

形时，首先将人脸区域划分为多种感兴趣区域，并且对多种感兴趣区域进行单独识别，然后结合多种感兴趣区域的识别结合进行融合分析，从而判断是否出现注意力不集中第二情形。这样，可以提高分析的准确性，对注意力不集中的识别效果较好。

[0022] 具体的，分别对人脸框区域以及人脸补充区域进行识别，可以识别出人员的多种注意力情况，例如驾驶员注视左后视镜、注视正前方、注视车内后视镜、注视右后视镜、注视仪表盘、注视中控区以及闭眼睛等情况，结合人脸框区域以及人脸补充区域的识别结果，可以判断驾驶员有无分心驾驶或者是否想变道行驶，当若干连续人脸图像被分成注视车内后视镜、注视仪表盘、注视中控区时可判断驾驶员分心驾驶，当若干连续人脸图像被分成注视左后视镜以及注视正前方时可判断驾驶员想变道行驶。

[0023] 而利用人脸框区域以及眼部区域的识别结果，可以判断驾驶员有无疲劳驾驶、分心驾驶或者是否想变道行驶，当连续若干人脸图像被分成闭眼睛时，可判断驾驶员疲劳驾驶，当若干连续人脸图像被分成偏左或者偏右，眼睛注视左侧或者右侧，可认为出现注意力不集中的第二情形。

[0024] 判断驾驶员是否出现注意力不集中的第一情形，只需要判断驾驶员的头部转动方向是否大于预设角度，例如向上、向下、向左或者向右转动超过 60° ，即可以认为出现注意力不集中的第一情形，这样，第一卷积神经网络的设计非常简单，运算量较小。一旦被判断为出现注意力不集中的第一情形，不需要执行第二情形的判断，可以节省注意力不集中判断的整体计算量。

[0025] 而第二卷积神经网络的计算量较为复杂，可以精确识别出驾驶员是否出现其他注意力不集中的情形，且判断更为准确。

[0026] 另外，由于在获取的视频数据中，连续帧的画面很相近，如果对每一帧都进行识别，将导致计算量非常巨大，而且进行大量的相似计算，计算结果也基本相同，因此，设定每连续多帧图像中，只选取其中一帧图像进行识别，可以大大减小注意力不集中识别的计算量，但又能够保证识别结果的准确性。优选的，可以从每六帧图像中选取一帧图像进行识别。

对附图的简要说明

附图说明

- [0027] 图1是本发明基于级联神经网络的注意力检测方法实施例的流程图。
- [0028] 图2是本发明基于级联神经网络的注意力检测方法实施例中数值的softmax概率值的计算公式。
- [0029] 图3是本发明基于级联神经网络的注意力检测方法实施例中第一卷积神经网络的结构框图。
- [0030] 图4是应用本发明基于级联神经网络的注意力检测方法实施例对图像进行识别的四种感兴趣区域的示意图。
- [0031] 图5是本发明基于级联神经网络的注意力检测方法实施例中第二卷积神经网络的结构框图。
- [0032] 图6是本发明基于级联神经网络的注意力检测方法实施例中第二卷积神经网络的瓶颈残差层步长为1时的结构框图。
- [0033] 图7是本发明基于级联神经网络的注意力检测方法实施例中第二卷积神经网络的瓶颈残差层步长为2时的结构框图。
- [0034] 以下结合附图及实施例对本发明作进一步说明。

发明实施例

本发明的实施方式

- [0035] 本发明的基于级联神经网络的注意力检测方法应用在智能设备上，优选的，智能设备设置有摄像装置，如摄像头等，智能设备利用摄像装置所获取的视频数据进行图像分析，进而判断特定的人员是否出现注意力不集中的情形。优选的，智能设备设置有处理器以及存储器，存储器上存储有计算机程序，处理器通过执行该计算机程序实现基于级联神经网络的注意力检测方法。
- [0036] 基于级联神经网络的注意力检测方法实施例：
- [0037] 本实施例主要是基于头部姿态和眼部信息，并应用级联卷积神经网络对特定人员的注意力进行，整个方法主要包括视频采集、图像处理以及注意力检测三个步骤。
- [0038] 在视频采集步骤中，利用摄像装置拍摄视频数据，本实施例可以针对不同场景（包括不同拍摄角度、外界光照条件、目标的位置等）的视频数据进行识别，

因此，摄像装置可以获得多种不同姿态下的目标视频数据。在图像处理步骤中，从视频数据中获取多帧图像，利用人脸检测算法检测帧，并截取人脸区域的图像。在注意力检测步骤中，先采用计算复杂度低的第一卷积神经网络判断检测对象的头部姿态，进而实现注意力检测的初级判断；然后，对检测到的感兴趣的人脸区域进一步进行截取和扩充，采用计算复杂度高的第二卷积神经网络提取头部姿态特征和眼部特征信息，通过分析人的注视方向，进而判断人的行为。本实施例所采用的级联卷积神经网络泛化性能好、计算复杂度低，而且适用于嵌入式设备。

[0039] 下面结合图1对本实施例的具体工作方法进行说明。首先，执行步骤S1，获取视频数据，即由智能设备的摄像装置获取连续的视频数据。具体的，智能设备可以是设置在车内的用于检测驾驶员注意力是否不集中的设备，摄像装置可以设置在驾驶位的正前方、侧前方等位置，例如设置在驾驶位遮阳板的下方、中控台上方等。摄像装置可以在汽车发动机启动后开始录制视频，并且将获取的连续的视频数据传输至处理器，由处理器对视频数据进行处理。

[0040] 然后，执行步骤S2，对图像进行识别，提取图像中的人脸区域。由于步骤S1获取的视频数据包括连续的多帧图像，步骤S2是对接收到的多帧图像进行识别。但由于连续的多帧图像的画面很相近，如果对每一帧图像均进行识别，不但导致计算量非常巨大，而且相邻的多帧图像的识别结果往往是相同的，因此，本实施例设定每连续的多帧图像中，可以选取其中一帧图像进行识别，例如设定从每六帧或者八帧图像中选取一帧图像进行识别，即对该帧图像进行人脸检测，并截取检测到的人脸区域。具体的，人脸检测的过程是假设在输入图像中存在一个或者多个人脸的情况下，确定图像中全部人脸的位置、大小和姿势的过程，该过程可以采用目前公知的人脸检测算法实现，在此不再赘述。

[0041] 然后，对提取的人脸区域进行注意力检测，具体的，执行步骤S3至步骤S10。本实施例通过一种级联的卷积神经网络来对图像进行识别，具体的，该级联的卷积神经网络包括第一神经网络以及第二神经网络，其中，第一卷积神经网络用于判断检测对象的头部姿态，进而实现注意力的初级判断，即判断是否出现注意力不集中的第一情形。第二卷积神经网络用于提取头部姿态特征、眼部特

征信息，通过分析人的注视方向，进而判断人的行为，实现注意力检测。

[0042] 具体的，先执行步骤S3，应用第一卷积神经网络对人脸区域进行识别。具体的，本实施例利用预先训练好的检测模型判断检测对象的注意力集中状态，例如可以设定检测对象的头部转动超过一定的角度时，属于注意力不集中的第一情形，如当驾驶员处于左转头（大于60度）、右转头（大于60度）、上抬头（大于60度）、下低头（大于60度）时均属于注意力不集中。

[0043] 因此，第一卷积神经网络的识别任务相对简单、容易区分，第一卷积神经网络是一个体积小、计算复杂度低的卷积神经网络。参见图3，本实施例的第一卷积神经网络包括若干卷积层(convolution)、若干池化层(maxpool)、一层全连接层16(fully connect)和一层分类层17，其中，每一个池化层位于相邻的两个卷积层之间，如图3所示的，虚线框11中包括多个卷积层与池化层组合而成的单元，每一个单元内包括一个卷积层以及一个池化层，在最后一个池化层的输出被输入至卷积层15，因此卷积层的数量比池化层的数量多1层。

[0044] 本实施例中，多个卷积层的参数有两种，其中一种卷积层的参数滤波器个数为 m 、卷积核大小为 $k_1 \times k_1$ 、步长像素为 S_1 ，另一种卷积层的参数是滤波器个数为 n 、卷积核大小为 $k_2 \times k_2$ 、步长像素为 S_2 。每一层池化层是对上一层卷积层的输出进行采样。全连接层16用于实现将卷积层15输出的二维特征矩阵变成一维特征向量的过程。分类层17作为第一卷积神经网络的最后一层，采用softmax函数，将多个神经元的输出映射到(0,1)区间内，其可以被理解成概率分布。假设概率分布向量为 P ， P_i 表示 P 中第 i 个数值，该数值的softmax概率值定义如图2的公式所示。

[0045] 在 P 中寻找最大值，将概率最大的 i 所对应的类别作为检测结果。该检测结果为驾驶员的头部转动角度是否超过预设的角度。

[0046] 然后，执行步骤S4，判断步骤S3的检测结果为是否驾驶员的头部转动角度超过预设角度，如是，则确认驾驶员出现注意力不集中的第一情形，此时执行步骤S9，发出警示信息，例如发出语音警示信息。

[0047] 如果确认没有出现注意力不集中的第一情形，则应用第二卷积神经网络判断是否出现注意力不集中的第二情形。具体的，先执行步骤S5，将人脸区域中截取

出多种感兴趣区域。参见图4，以驾驶员坐在在驾驶位置为例，第一种感兴趣区域是直接得到嵌入式人脸视野，其对应的人的图像信息是后视镜和驾驶员左侧镜子中间部分，虚线框21内的图像部分，第一感兴趣区域无需进行人脸检测操作，可直接利用该图像信息判断人的注意力。第二种感兴趣区域是利用已知的人脸检测算法检测并截取人脸框作为输入图像，实线框22内的图像区域，第二感兴趣区域可以被称为人脸框区域。第三种感兴趣区域是在第二种感兴趣区域的基础之上，对检测出的人脸框沿着上下左右四个方向进行扩充，添加了人脸额外部分信息，如图中实线框23内的图像区域，第三种感兴趣区域可以被称为人脸补充区域。第三种感兴趣区域的截取方法添加了额外的辅助特征，不仅能够确定人的头部位置而且具有很好的鲁棒性。第四种感兴趣区域是在第二种感兴趣区域的基础上只截取人脸的上半部分，如图4的实线框24内的图像区域，因此第四种感兴趣区域是眼部区域，主要用于通过关注眼部信息来判断驾驶员的注意力情况。

[0048] 然后，执行步骤S6，应用第二卷积神经网络对多种感兴趣区域进行识别。例如，利用预先训练好的注意力检测模型对多种感兴趣区域进行识别并对识别结果进行分类。以驾驶员坐在驾驶位置为例，检测对象的注意力分为七类，分别是注视左后视镜、注视正前方、注视车内后视镜、注视右后视镜、注视仪表盘、注视中控区以及闭眼睛。在其它应用场景中，可以将检测对象的注意力分为六类，分别是注视左侧、注视右侧、注视正前方、注视上方、注视下方以及闭眼睛。

[0049] 由于步骤S6的识别、分类任务比较复杂，尤其相连区域如正前方和仪表盘区分难度比较大，因此，本实施例使用一个学习能力强、运算速度快的卷积神经网络进行识别，即应用第二卷积神经网络实现上述识别工作。参见图5，本实施例的第二卷积神经网络包括依次级联的第一卷积层31(convolution)、逐深度卷积层32 (depthwise convolution)、若干瓶颈残差层(bottleneck residual)、第二卷积层35、线性全局逐深度卷积层36(linear GDConv)、线性卷积层37(linear Conv)、全连接层38(fully connect)以及分类层39。其中，图5的虚线框表示多个瓶颈残差层组成的单元，例如多个瓶颈残差层包括瓶颈残差层33、34等，对于第i个瓶颈残差

层，瓶颈残差层重复 n_i 次，每一层通道扩充数为 t_i ，步长为 s_i 。

[0050] 本实施例中，第一卷积层31、第二卷积层35的参数可以是不相同的，其中一个卷积层的参数是滤波器个数为 m 、卷积核大小为 $k_1 \times k_1$ 、步长像素为 S_1 ，另一个卷积层的参数是滤波器个数为 n 、卷积核大小为 $k_2 \times k_2$ 、步长像素为 S_2 。

[0051] 逐深度卷积层32是对输入的每一个通道单独用对应通道的卷积核进行卷积操作，假设输入维度是 m ，尺寸是 $w \times h$ ，对应卷积层滤波器个数是 m ，卷积核大小是 $k \times k$ ，采用逐深度卷积操作，此时输出维度为 m ，尺寸大小为 $w' \times h'$ 。

[0052] 每一个瓶颈残差层包含卷积单元、逐深度卷积单元和残差单元，其中，逐深度卷积单元用于接收卷积单元的输出，残差单元在卷积单元的步长为1时，实现瓶颈残差层的残差运算。如图6所示，残差单元在卷积单元的步长为1、通道数为 c' 时，此时输入和输出的对应通道上的值相加实现残差运算，即输入的数据经过依次级联的第一卷积单元41、逐深度卷积单元42、第二卷积单元43以及残差单元44，在残差单元44上实现输入与输出的累加计算。

[0053] 残差单元在卷积单元的步长为2、通道数为 c' 时，此时输入的维度为 $[w, h]$ ，输出的维度为 $[w', h']$ ，由于输入的维度与输出的维度不相等，此时不进行残差运算，此时的结构框图如图7所示，输入依次经过第一卷积单元51、逐深度卷积单元52以及第二卷积单元53后输出。

[0054] 线性全局逐深度卷积层36的卷积核大小与输入尺寸的大小相同，滤波器个数为 m ，卷积核大小为 $k \times k$ ，此时输入通道数为 n ，尺寸大小也为 $k \times k$ ，经过线性全局逐深度卷积运算，输出通道数为 m ，尺寸大小为 1×1 。

[0055] 线性卷积层37是一种特殊形式的卷积层，其采用线性函数作为该层的激活函数。全连接层38的计算过程是将上层输出的二维特征矩阵变成一维特征向量的过程，输出的维度与分类个数相同。分类层39的计算方法与第一卷积神经网络的分类层17的计算方法相同，在此不再赘述。

[0056] 然后，执行步骤S7，根据步骤S6分别对四种感兴趣区域的识别结果进行融合分析，得到融合分析的结果。具体的，以驾驶员坐在驾驶位置为例，利用第二种感兴趣区域以及第三种感兴趣区域的识别结果进行融合分析时，首先采用人脸检测算法检测出人脸区域，并截取对应的人脸框图像，实现人脸框的分类；然

后对截取的人脸框沿着四个方向上的扩充得到新图像，实现新图像的分类，利用二者分类的结果可以判断驾驶员有无分心驾驶或者是否想变道行驶，当若干连续人脸图像被分成注视车内后视镜、注视仪表盘、注视中控区时可判断驾驶员分心驾驶，当若干连续人脸图像被分成注视左后视镜以及注视正前方时可判断驾驶员想变道行驶。

[0057] 又例如，利用第二种感兴趣区域以及第四种感兴趣区域的识别结果进行融合分析时，先采用人脸检测算法检测出人脸区域，并截取对应的人脸框图像，实现人脸框的分类，然后保留人脸框的上半部分，即获取眼部区域的信息，同时对眼部信息进行分类，利用二者的分类结果进行融合分析，可以判断驾驶员有无疲劳驾驶、分心驾驶或者是否想变道行驶，当连续若干人脸图像被分成闭眼睛时，可判断驾驶员疲劳驾驶。可选的，该方法还可应用在其他场景，例如学生课堂注意力的检测，结合第二种感兴趣区域以及第四种感兴趣区域的识别结果，当若干连续人脸图像被分成偏左或者偏右、眼睛注视左侧或者右侧，可认为是注意力不集中的情形。

[0058] 然后，执行步骤S7，根据步骤S6的分析结果，判断是否出现注意力不集中的第二情形，例如出现疲劳驾驶、分心驾驶等情况，如是，则执行步骤S9，发出警示信息，否则，执行步骤S10，根据步骤S7的分析结果对驾驶员的行为进行预测，例如希望向左变道行驶等，该预测结果可以提供给其他算法使用。例如，在辅助驾驶领域，可以根据步骤S7的结果，判断驾驶员希望向左变道行驶时，可以检测左侧后方来车情况，如左侧后方一定距离内是否有行驶的车辆等，从而给驾驶员发出指示信息。

[0059] 可选的，第二卷积神经网络可以被替换为计算能力强更加轻量的网络架构，如 ShuffleNet，也可以减少卷积神经网络的瓶颈残差层，重新对该模型进行训练。

[0060] 计算机装置实施例：

[0061] 本实施例的计算机装置可以是智能设备，例如具有图像处理能力的车载监视仪器等，该计算机装置包括有处理器、存储器以及存储在存储器中并可在处理器上运行的计算机程序，处理器执行计算机程序时实现上述基于级联神经网络的

注意力检测方法的各个步骤。

[0062] 例如，计算机程序可以被分割成一个或多个模块，一个或者多个模块被存储在存储器中，并由处理器执行，以完成本发明的各个模块。一个或多个模块可以是能够完成特定功能的一系列计算机程序指令段，该指令段用于描述计算机程序在终端设备中的执行过程。

[0063] 本发明所称处理器可以是中央处理单元(Central Processing Unit, CPU)，还可以是其他通用处理器、数字信号处理器 (Digital Signal Processor, DSP)、专用集成电路 (Application Specific Integrated Circuit, ASIC)、现成可编程门阵列 (Field-Programmable Gate Array, FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等，处理器是终端设备的控制中心，利用各种接口和线路连接整个终端设备的各个部分。

[0064] 存储器可用于存储计算机程序和/或模块，处理器通过运行或执行存储在存储器内的计算机程序和/或模块，以及调用存储在存储器内的数据，实现终端设备的各种功能。存储器可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）等；存储数据区可存储根据手机的使用所创建的数据（比如音频数据、电话本等）等。此外，存储器可以包括高速随机存取存储器，还可以包括非易失性存储器，例如硬盘、内存、插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0065] 计算机可读存储介质：

[0066] 上述计算机装置所存储的计算机程序如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明实现上述实施例方法中的全部或部分流程，也可以通过计算机程序来指令相关的硬件来完成，该计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器执行时，可实现上述基于级联神经网络的注意力检测方法的各个步骤。

[0067] 其中，计算机程序包括计算机程序代码，计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。计算机可读介质可以包括：能够携带计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、电载波信号、电信信号以及软件分发介质等。需要说明的是，计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减，例如在某些司法管辖区，根据立法和专利实践，计算机可读介质不包括电载波信号和电信信号。

[0068] 最后需要强调的是，本发明不限于上述实施方式，如对多种感兴趣区域划分方法的变化，或者根据多种感兴趣区域识别结果融合分析具体过程及结果的改变等，这些改变也应该包括在本发明权利要求的保护范围内。

[0069] 工业应用性

[0070] 本发明的方法使用级联的卷积神经网络进行识别，由于第一级卷积神经网络计算复杂度较低，可以对简单的场景进行分析，并判断出驾驶员是否出现注意力不集中的第一情形，这样可以减小整个卷积神经网络的运算量，整个卷积神经网络模型体积小、计算复杂度低。

[0071] 此外，本发明的方法先利用头部姿态信息初步判断是否注意力集中，然后利用头部姿态和眼部信息进一步来检测驾驶员的注意力；在检测驾驶员的注意力之前，采用四种方法对原始图像进行处理得到四种感兴趣区域，将分类结果进行融合，分析人的行为与意图，因此，本发明的级联卷积神经网络泛化性能好、计算复杂度低，而且适用于嵌入式设备。

权利要求书

- [权利要求 1] 基于级联神经网络的注意力检测方法，包括：
获取视频数据，并对多帧图像进行识别，提取多帧图像的人脸区域；
其特征在于：
应用第一卷积神经网络对所述人脸区域进行识别，判断是否出现注意力不集中的第一情形；
如确认没有出现注意力不集中的第一情形，则应用第二卷积神经网络对所述人脸区域进行识别，判断是否出现注意力不集中的第二情形；
其中，所述第一卷积神经网络的计算复杂度小于所述第二卷积神经网络的计算复杂度。
- [权利要求 2] 根据权利要求1所述的基于级联神经网络的注意力检测方法，其特征在于：
应用所述第二卷积神经网络对所述人脸区域进行识别包括：从所述人脸区域划截取出多种感兴趣区域，根据二种以上的所述感兴趣区域的识别结果判断是否出现注意力不集中的第二情形。
- [权利要求 3] 根据权利要求2所述的基于级联神经网络的注意力检测方法，其特征在于：
多种所述感兴趣区域包括人脸框区域以及人脸补充区域；
根据二种以上的所述感兴趣区域的识别结果判断是否出现注意力不集中的第二情形包括：根据所述人脸框区域以及所述人脸补充区域的图像识别结果判断是否出现注意力不集中的第二情形。
- [权利要求 4] 根据权利要求2所述的基于级联神经网络的注意力检测方法，其特征在于：
多种所述感兴趣区域包括人脸框区域以及眼部区域；
根据二种以上的所述感兴趣区域的识别结果判断是否出现注意力不集中的第二情形包括：根据所述人脸框区域以及所述眼部区域的图像识别结果判断是否出现注意力不集中的第二情形。
- [权利要求 5] 根据权利要求1至4任一项所述的基于级联神经网络的注意力检测方法，

其特征在于：

判断是否出现注意力不集中的第一情形包括：应用第一卷积神经网络对所述人脸区域进行识别，判断头部向预设方向的转动角度是否大于预设角度，如是，确认出现注意力不集中的第一情形。

[权利要求 6] 根据权利要求1至4任一项所述的基于级联神经网络的注意力检测方法，其特征在于：

所述第二卷积神经网络包括依次级联的第一卷积层、逐深度卷积层、多个瓶颈残差层、第二卷积层、线性全局逐深度卷积层、线性卷积层、全连接层以及分类层。

[权利要求 7] 根据权利要求6所述的基于级联神经网络的注意力检测方法，其特征在于：

所述瓶颈残差层包括卷积单元以及接收所述卷积单元输出的逐深度卷积单元，还设置有残差单元，所述残差单元在所述卷积单元的步长为1时，实现所述瓶颈残差层的残差运算。

[权利要求 8] 根据权利要求1至4任一项所述的基于级联神经网络的注意力检测方法，其特征在于：

获取所述视频数据后对多帧所述图像进行识别包括：从所述视频数据的每连续预设帧数的图像中选取一帧所述图像进行识别。

[权利要求 9] 计算机装置，其特征在于，包括处理器以及存储器，所述存储器存储有计算机程序，所述计算机程序被处理器执行时实现如权利要求1至8中任意一项所述基于级联神经网络的注意力检测方法的各个步骤。

[权利要求 10] 计算机可读存储介质，其上存储有计算机程序，其特征在于：所述计算机程序被处理器执行时实现如权利要求1至8中任意一项所述基于级联神经网络的注意力检测方法的各个步骤。

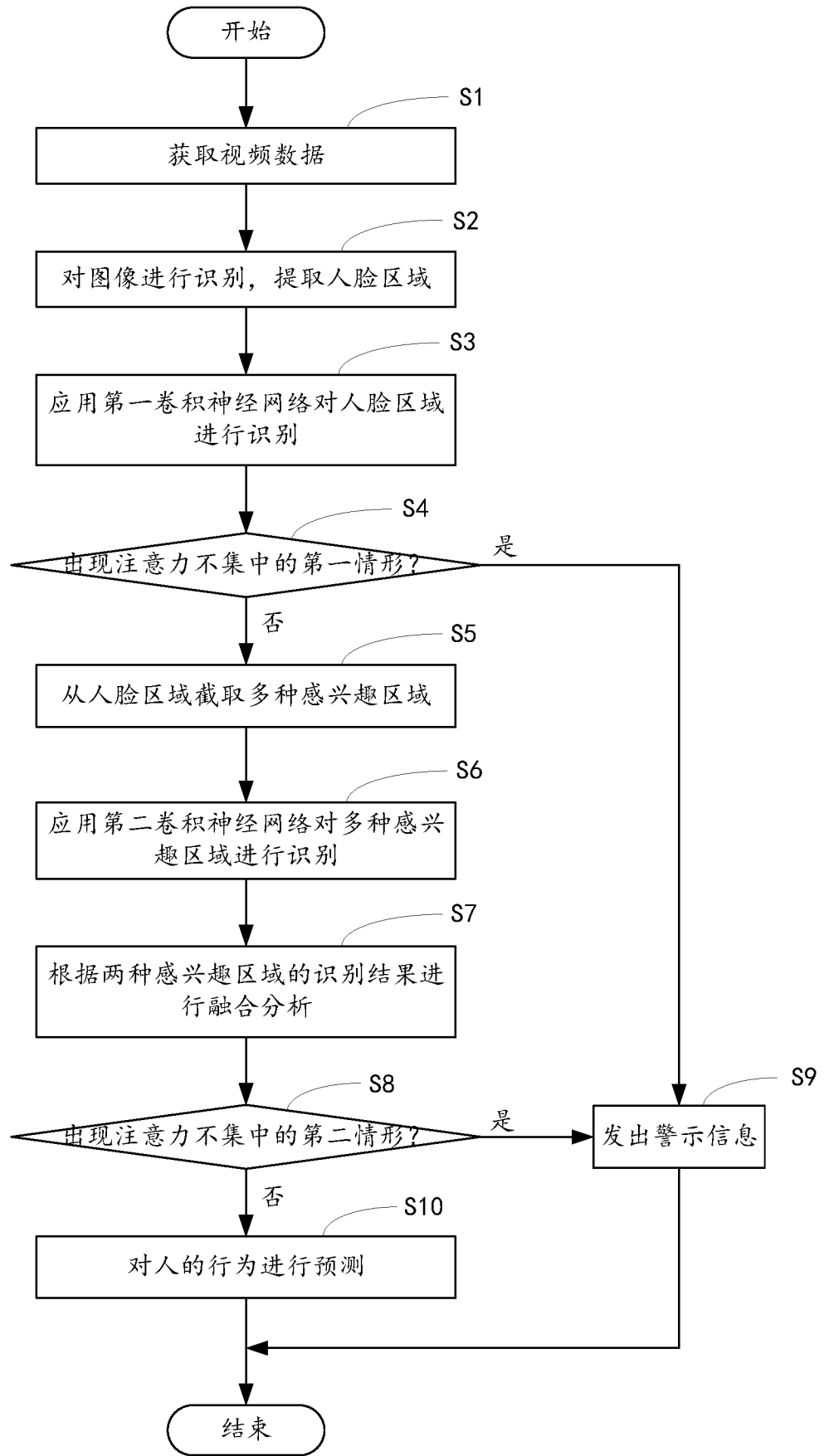


图 1

$$P_i = \frac{e^i}{\sum_j e^j}, i = 1, 2, \dots, n$$

图 2

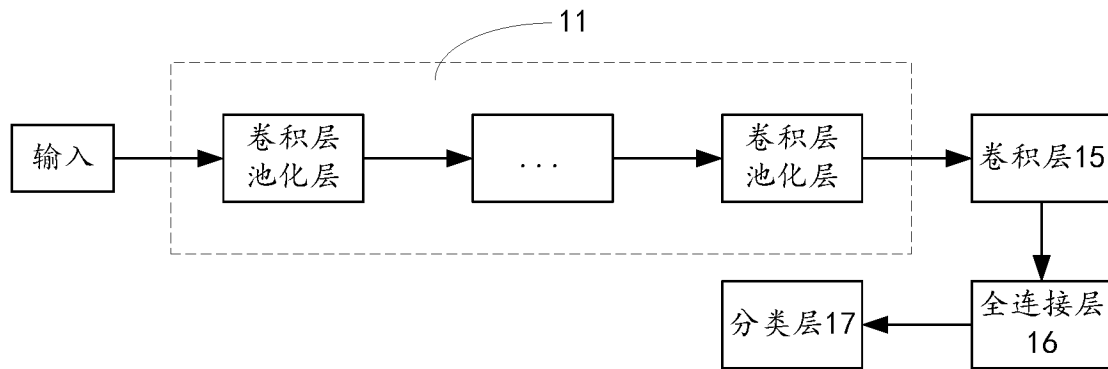


图 3

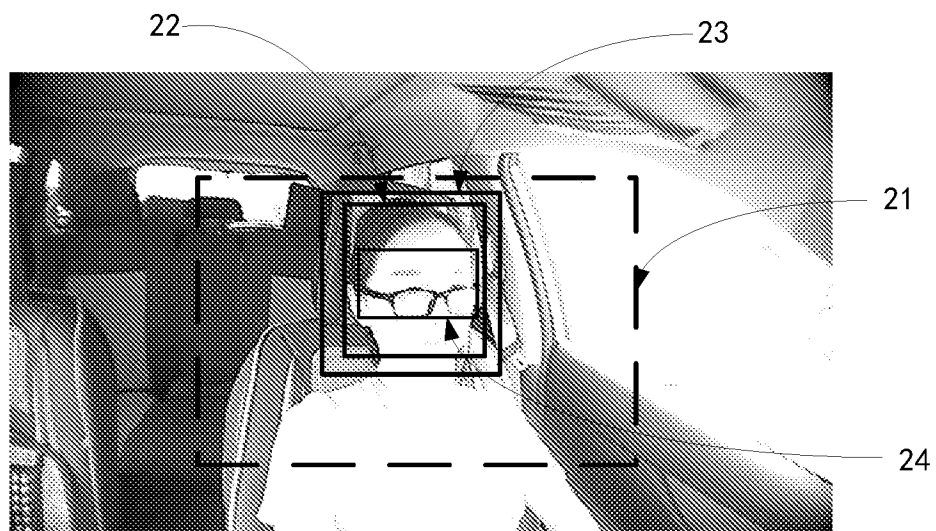


图 4

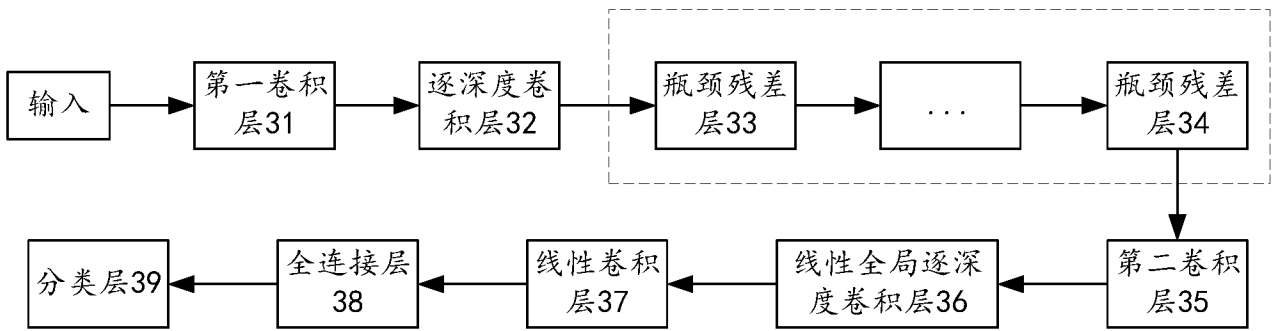


图 5

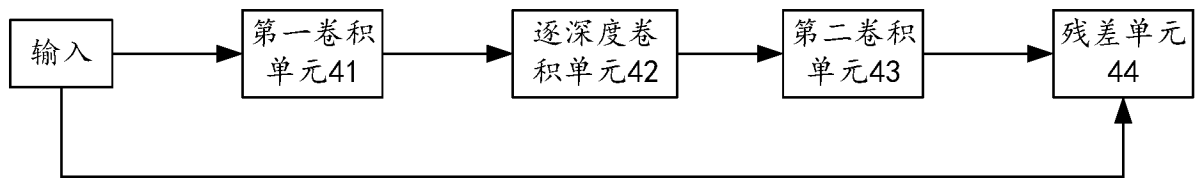


图 6

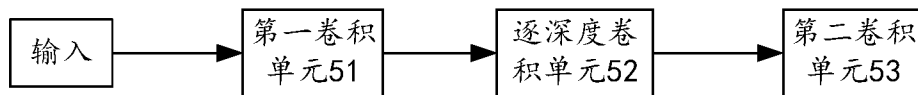


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/098407

A. CLASSIFICATION OF SUBJECT MATTER		
G06K 9/00(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
G06K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
CNTXT; CNABS; DWPI; CNKI; SIPOABS; 人脸, 面部, 头部, 识别, 图像, 视频, 疲劳, 驾驶, 注意力, 卷积, 神经网络, 感兴趣区域, 眼眶, 眼球, 眼部; human, face, head, recognition, image, fatigue, drive, attention, convolve, neural network, ROI, region of interest, eye socket, eyeball.		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 109740477 A (DIAS AUTOMOTIVE ELECTRONIC SYSTEMS CO., LTD.) 10 May 2019 (2019-05-10) description, paragraphs [0063]-[0099]	1-10
Y	CN 109598174 A (XIAMEN CLARION ELECTRICAL ENTERPRISE CO., LTD.) 09 April 2019 (2019-04-09) description, paragraphs [0086]-[0230]	1-10
A	CN 108664947 A (WUYI UNIVERSITY) 16 October 2018 (2018-10-16) entire document	1-10
A	JP 4967559 B2 (TOYOTA CENTRAL R&D LABS., INC.) 04 July 2012 (2012-07-04) entire document	1-10
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
11 April 2020		23 April 2020
Name and mailing address of the ISA/CN		Authorized officer
China National Intellectual Property Administration (ISA/ CN) No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088 China		
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/098407

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109740477	A	10 May 2019	None			
CN	109598174	A	09 April 2019	None			
CN	108664947	A	16 October 2018	None			
JP	4967559	B2	04 July 2012	JP	2008077189	A	03 April 2008

国际检索报告

国际申请号

PCT/CN2019/098407

<p>A. 主题的分类</p> <p>G06K 9/00 (2006.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																	
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06K</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNXTX;CNABS;DWPI;CNKI;SIPOABS: 人脸, 面部, 头部, 识别, 图像, 视频, 疲劳, 驾驶, 注意力, 卷积, 神经网络, 感兴趣区域, 眼眶, 眼球, 眼部; human, face, head, recognition, image, fatigue, drive, attention, convolve, neural network, ROI, region of interest, eye socket, eyeball.</p>																	
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 109740477 A (联创汽车电子有限公司) 2019年 5月 10日 (2019 - 05 - 10) 说明书第[0063]-[0099]段</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>CN 109598174 A (厦门歌乐电子企业有限公司) 2019年 4月 9日 (2019 - 04 - 09) 说明书第[0086]-[0230]段</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>CN 108664947 A (五邑大学) 2018年 10月 16日 (2018 - 10 - 16) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>JP 4967559 B2 (株式会社豊田中央研究所) 2012年 7月 4日 (2012 - 07 - 04) 全文</td> <td>1-10</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 109740477 A (联创汽车电子有限公司) 2019年 5月 10日 (2019 - 05 - 10) 说明书第[0063]-[0099]段	1-10	Y	CN 109598174 A (厦门歌乐电子企业有限公司) 2019年 4月 9日 (2019 - 04 - 09) 说明书第[0086]-[0230]段	1-10	A	CN 108664947 A (五邑大学) 2018年 10月 16日 (2018 - 10 - 16) 全文	1-10	A	JP 4967559 B2 (株式会社豊田中央研究所) 2012年 7月 4日 (2012 - 07 - 04) 全文	1-10
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
Y	CN 109740477 A (联创汽车电子有限公司) 2019年 5月 10日 (2019 - 05 - 10) 说明书第[0063]-[0099]段	1-10															
Y	CN 109598174 A (厦门歌乐电子企业有限公司) 2019年 4月 9日 (2019 - 04 - 09) 说明书第[0086]-[0230]段	1-10															
A	CN 108664947 A (五邑大学) 2018年 10月 16日 (2018 - 10 - 16) 全文	1-10															
A	JP 4967559 B2 (株式会社豊田中央研究所) 2012年 7月 4日 (2012 - 07 - 04) 全文	1-10															
<input type="checkbox"/> 其余文件在C栏的续页中列出。		<input checked="" type="checkbox"/> 见同族专利附件。															
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p>		<p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p>															
<p>国际检索实际完成的日期</p> <p>2020年 4月 11日</p>		<p>国际检索报告邮寄日期</p> <p>2020年 4月 23日</p>															
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>受权官员</p> <p>钟阳雪</p> <p>电话号码 86-010-62411634</p>															

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/098407

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	109740477	A	2019年 5月 10日	无	
CN	109598174	A	2019年 4月 9日	无	
CN	108664947	A	2018年 10月 16日	无	
JP	4967559	B2	2012年 7月 4日	JP 2008077189	A 2008年 4月 3日