

(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
2022 年 11 月 17 日 (17.11.2022)



(10) 国际公布号
WO 2022/237481 A1

- (51) 国际专利分类号:
G06K 9/00 (2022.01) *G06K 9/62* (2022.01)
- (21) 国际申请号: PCT/CN2022/088032
- (22) 国际申请日: 2022 年 4 月 20 日 (20.04.2022)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110519356.6 2021年5月12日 (12.05.2021) CN
- (71) 申请人: 北京百度网讯科技有限公司 (BEIJING BAIDU NETCOM SCIENCE TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市海淀区上地十街10号百度大厦二层, Beijing 100085 (CN)。
- (72) 发明人: 卢子鹏 (LU, Zipeng); 中国北京市海淀区上地十街10号百度大厦二层, Beijing 100085 (CN)。王健 (WANG, Jian); 中国北京市海淀区上地十街10号百度大厦二层, Beijing 100085 (CN)。孙昊 (SUN, Hao); 中国北京市海淀区上地十街10号百度大厦二层, Beijing 100085 (CN)。丁二锐 (DING, Errui); 中国北京市海淀区上地十街10号百度大厦二层, Beijing 100085 (CN)。
- (74) 代理人: 北京清亦华知识产权代理事务所 (普通合伙) (TSINGYIHUA INTELLECTUAL PROPERTY LLC); 中国北京市海淀区北洼路45号1号楼2层201, Beijing 100142 (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,

(54) **Title:** HAND-RAISING RECOGNITION METHOD AND APPARATUS, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 举手识别方法、装置、电子设备及存储介质

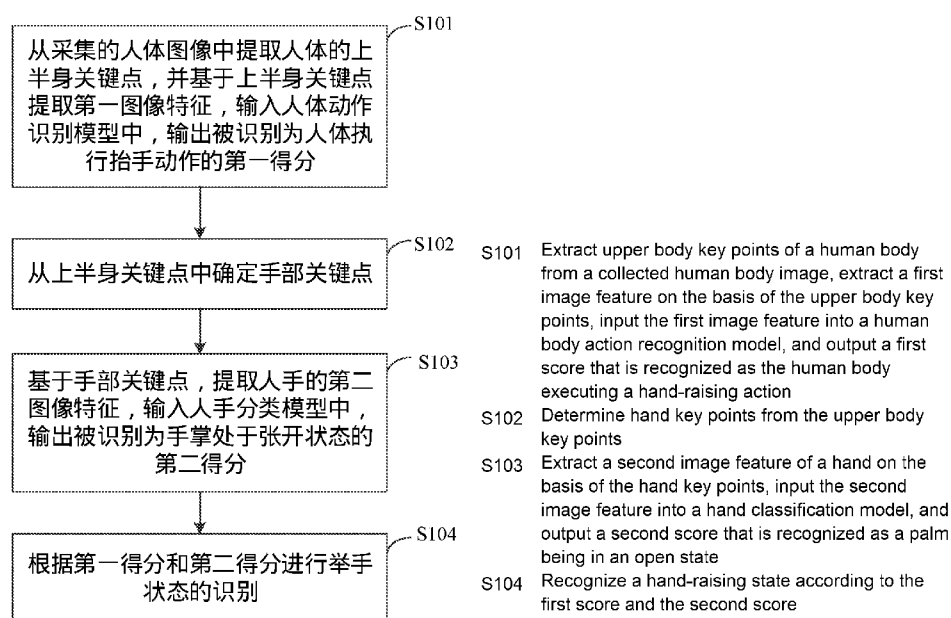


图 1

(57) **Abstract:** A hand-raising recognition method and apparatus, an electronic device, and a storage medium. The method comprises: extracting upper body key points of a human body from a collected human body image, extracting a first image feature on the basis of the upper body key points, inputting the first image feature into a human body action recognition model, and outputting a first score that is recognized as the human body executing a hand-raising action; determining hand key points from the upper body key points; extracting a second image feature of a hand on the basis of the hand key points, inputting the second image feature into a hand classification model, and outputting a second score that is recognized as a palm being in an open state; and recognizing a hand-raising state according to the first score and the second score.



GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, 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)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种举手识别方法、装置、电子设备及存储介质, 方法包括: 从采集的人体图像中提取人体的上半身关键点, 并基于上半身关键点提取第一图像特征, 输入人体动作识别模型中, 输出被识别为人体执行抬手动作的第二得分; 从上半身关键点中确定手部关键点; 基于手部关键点, 提取人手的第二图像特征, 输入人手分类模型中, 输出被识别为手掌处于张开状态的第三得分; 根据第二得分和第三得分进行举手状态的识别。

举手识别方法、装置、电子设备及存储介质

5 相关申请的交叉引用

本公开基于申请号为 202110519356.6、申请日为 2021 年 05 月 12 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此引入本公开作为参考。

技术领域

10 本公开涉及图像处理领域，尤其涉及计算机视觉和深度学习等人工智能领域。

背景技术

随着技术的发展，很多场景存在举手识别的需求，比如课堂上对于学生的举手识别。相关技术中，通常是将采集到的待识别图像直接输入训练检测模型，针对待识别图像直接进行图像识别。

15 一般情况下，由于图像采集设备的布放位置，所获取到的待识别图像为远景图像，其中，人手部位通常较小，且不够清晰，人体举手动作特征相对模糊，使得举手识别的准确率较低。因此，如何在清晰度不够的远景图像中进行高准确率的举手识别成为了需要解决的问题。

发明内容

20 本公开提出了一种用于举手识别方法、装置、电子设备以及存储介质、计算机程序产品。

根据本公开的一方面，提出了一种举手识别方法，包括：从采集的人体图像中提取人体的上半身关键点，并基于所述上半身关键点提取第一图像特征，输入人体动作识别模型中，输出被识别为人体执行抬手动作的第一得分；从所述上半身关键点中确定手部关键点；基于所述手部关键点，提取人手的第二图像特征，输入人手分类模型中，输出被识别为手掌处于张开状态的第二得分；根据所述第一得分和所述第二得分进行举手状态的识别。

25 根据本公开的第二方面，提出了一种举手识别装置，包括：提取模块，用于从采集的人体图像中提取人体的上半身关键点，并基于所述上半身关键点提取第一图像特征，输入人体动作识别模型中，输出被识别为人体执行抬手动作的第一得分；所述提取模块，还用于从所述上半身关键点中确定手部关键点；所述提取模块，还用于基于所述手部关键点，提取人手的第二图像特征，输入人手分类模型中，输出被识别为手掌处于张开状态的第二得分；识别模块，用于根据所述第一得分和所述第二得分进行举手状态的识别。

根据本公开的第三方面，提出了一种电子设备，包括：至少一个处理器；以及与所述至少一个处理器通信连接的存储器；其中，所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行上述第一方面提出的举手识别方法。

35 根据本公开的第四方面，提出了一种存储有计算机指令的非瞬时计算机可读存储介质，其中，所述计算机指令用于使所述计算机执行上述第一方面提出的举手识别方法。

根据本公开的第五方面，提出了一种计算机程序产品，包括计算机程序，所述计算机程序在被处理器执行时实现上述第一方面提出的举手识别方法。

应当理解，本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征，也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

附图用于更好地理解本方案，不构成对本公开的限定。其中：

45 图 1 是本公开一实施例的举手识别方法的流程示意图；
图 2 是本公开另一实施例的举手识别方法的流程示意图；
图 3 是本公开另一实施例的举手识别方法的流程示意图；

图 4 是本公开另一实施例的举手识别方法的流程示意图；
图 5 是本公开另一实施例的举手识别方法的流程示意图；
图 6 是本公开另一实施例的举手识别方法的流程示意图；
图 7 是本公开另一实施例的举手识别方法的流程示意图；
图 8 是本公开一实施例的举手识别装置的结构示意图；
图 9 是本公开另一实施例的举手识别装置的结构示意图；
图 10 是本公开一实施例的电子设备的示意性框图。

具体实施方式

以下结合附图对本公开的示范性实施例做出说明，其中包括本公开实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本公开的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

图像处理（Image Processing），用计算机对图像进行分析，以达到所需结果的技术。又称影像处理。图像处理一般指数字图像处理。数字图像是指用工业相机、摄像机、扫描仪等设备经过拍摄得到的一个大的二维数组，该数组的元素称为像素，其值称为灰度值。图像处理技术一般包括图像压缩，增强和复原，匹配、描述和识别 3 个部分。

深度学习（Deep Learning，简称 DL），是机器学习（Machine Learning，简称 ML）领域中一个新的研究方向，它被引入机器学习使其更接近于最初的目标——人工智能。深度学习是学习样本数据的内在规律和表示层次，这些学习过程中获得的信息对诸如文字，图像和声音等数据的解释有很大的帮助。它的最终目标是让机器能够像人一样具有分析学习能力，能够识别文字、图像和声音等数据。深度学习是一个复杂的机器学习算法，在语音和图像识别方面取得的效果，远远超过先前相关技术。

计算机视觉（Computer Vision），是一门研究如何使机器“看”的科学，更进一步的说，就是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科，计算机视觉研究相关的理论和技术，试图建立能够从图像或者多维数据中获取“信息”的人工智能系统。这里所指的信息指 Shannon 定义的，可以用来帮助做一个“决定”的信息。因为感知可以看作是从感官信号中提取信息，所以计算机视觉也可以看作是研究如何使人工系统从图像或多维数据中“感知”的科学。

人工智能（Artificial Intelligence，简称 AI），是研究使计算机来模拟人生的某些思维过程和智能行为（如学习、推理、思考、规划等）的学科，既有硬件层面的技术，也有软件层面的技术。人工智能硬件技术一般包括计算机视觉技术、语音识别技术、自然语言处理技术以及及其学习/深度学习、大数据处理技术、知识图谱技术等几大方面。

手势识别（Gesture recognition）手势识别在计算机科学中，是通过数学算法来识别人类手势的一个议题。手势识别可以来自人的身体各部位的运动，但一般是指脸部和手的运动。用户可以使用简单的手势来控制或与设备交互，让计算机理解人类的行为。其核心技术为手势分割、手势分析以及手势识别。手势识别可以被视为计算机理解人体语言的方式，从而在机器和人之间搭建比原始文本用户界面或甚至图形用户界面更丰富的桥梁。

智慧教育，即教育信息化，是指在教育领域（教育管理、教育教学和教育科研）全面深入地运用现代信息技术来促进教育改革与发展的过程。其技术特点是数字化、网络化、智能化和多媒体化，基本特征是开放、共享、交互、协作、泛在。以教育信息化促进教育现代化，用信息技术改变传统模式。

图 1 是本公开一实施例的举手识别方法的流程示意图，如图 1 所示，该方法包括：

S101，从采集的人体图像中提取人体的上半身关键点，并基于上半身关键点提取第一图像特征，输入人体动作识别模型中，输出被识别为人体执行抬手动作的的第一得分。

人体的运动动作执行依赖于骨骼、骨骼肌、骨关节等多种人体结构，针对人体的上肢，其基

本运动形式可以为推、拉等基础动作。在每个动作的运动过程中，骨关节为与其相连接的骨骼的转动提供了支点，因此，在本公开实施例中，可以将每个骨关节视为一个关键点，连接骨关节与骨关节之间的骨骼、骨骼肌等人体结构可以视为关键点之间的连接，进而使得可以基于提取到的关键点之间相对位置的变化实现人体动作的识别。

5 一般情况下，举手识别可以用于手势识别和智慧教育场景下，比如上课场景、投票决议场景等等。可以在场景中布放图像采集装置，依托于图像采集装置获取待识别的人体图像，进而实现对图像中的人体的上肢关键点的采集。

可选地，可以通过卷积神经网络（Convolutional Neural Networks，CNN）等算法进行图像中人体的上肢关键点的采集。

10 本公开实施例中，采集到人体上半身的关键点后，可以将提取到的关键点基于人体的基本结构进行连接，进而生成被采集的人体的上半身的图像特征，即为第一图像特征，其中，第一图像特征可以是基于人体上半身的关键点生成的拓扑结构图像特征。

其中，提取到的人体上半身的关键点，可以包括人体的头部、鼻子、颈、左眼、左耳、右眼、右耳、左肩、左肘、左腕、右肩、右肘、右腕、胸、左髋、右髋等等。

15 进一步地，如图2所示，图2为人体关键点的分布示意图，设定针对采集图像中的人体提取到图2中的人体关键点，其中，关键点1为人体上半身的左肩关节，关键点2为人体左胳膊上的肘关节，关键点3为左手腕关节，将关键点1、关键点2以及关键点3之间基于人体的骨骼结构相连接，即为关键点1与关键点2之间相连接，关键点2与关键点3之间相连接，关键点1与关键点3之间没有直接连接。进一步地，可以生成关键点1、关键点2以及关键点3生成采集图像中的人体左臂的图像特征。

20 可以设置一个人体动作识别模型，该模型可以基于第一图像特征实现人体抬手动作的识别，比如时空图卷积网络模型（Spatial Temporal Graph Convolutional Networks for Skeleton Based Action Reco，gcn），该模型基于人体的骨骼图序列制定，每一个关键点存在对应的人体的一个关节。

实现中，可以对人体动作识别模型进行一个初始的架构，进而根据输入的人体关键点的拓扑结构特征进行人体动作识别模型的训练，将训练样本人体关键点的拓扑结构特征输入至人体动作识别模型中，通过其特征提取层获取人体关键点的拓扑结构特征中携带的人体姿态信息，比如手臂抬起幅度等等，进而实现对人体关键点的拓扑结构特征中的人体所执行的动作的识别。通过多轮次的迭代，进而获取到成熟且精准的人体动作识别模型。

25 进一步地，将获取到的第一图像特征输入至成熟的人体动作识别模型中，通过模型的输出结果判断第一图像特征所属的采集图像中的人体是否执行了抬手动作。其中，人体动作识别模型输出的结果为分数，输入第一图像特征后获取到的输出结果即为本公开实施例中的第一得分。

S102，从上半身关键点中确定手部关键点。

30 基于采集的人体图像中提取到的上半身的关键点，包括人体上半身的全部关节以及其它可以体现人体运动动作的部位和/或组织结构。可以根据提取到的关键点之间的相对位置，可以确定人体手部位置的关键点。

35 本公开实施例中，人体手部的关键点可以包括人体的肘关节对应的关键点、腕关节对应的关键点、指关节对应的关键点等等，并且，可以将肘关节对应的关键点作为本公开实施例中的人体手部关键点与其他组织结构部位对应的关键点之间的分界。

40 一般情况下，人体手部的关键点中，肘关节对应的关键点与肩关节对应的关键点之间存在直接连接的关系，因此，可以将人体手部的关键点视为一个整体，通过该整体关键点与人体上半身其他组织结构对应的关键点之间的相对位置，从全部的上半身关键点中确定人体的手部关键点。

S103，基于手部关键点，提取人手的第二图像特征，输入人手分类模型中，输出被识别为手掌处于张开状态的第二得分。

45 人体的上肢的动作中，胳膊抬起的动作可以是举手也可以是其他动作，一般情况下，在教育场景、投票场景等需要人体执行举手动作的场景，人体除了手臂抬起之外，手掌也会处于张开状态。基于该设定，在人体手臂抬起的上肢动作中，需要进行进一步的筛选，避免人体执行抬手动

作且手掌处于非张开状态的动作的噪声干扰，进而提高人体举手识别的准确率。

基于获取到的手部关键点，将手部关键点与关键点之间基于人体的基础结构相连接。比如，肘关节对应的关键点需要与腕关节对应的关键点沿手肘至手腕方向相连接，腕关节对应的关键点需要与距离其最近的指关节对应的关键点沿手腕至手指方向相连接。

5 进一步地，基于手部关键点的位置以及关键点之间的连接关系，生成人体的手部的图像特征，即为第二图像特征。将第二图像特征输入预设的人体手部状态的识别模型，获取到模型的输出结果，并基于结果判断人体的手部是否处于张开状态。

其中，人体的手部状态识别模型的输出结果为分数，输入第二图像特征后获取到的即为第二图像特征所属的采集图像中的人体手掌是否处于张开状态的第二得分。

10 S104，根据第一得分和第二得分进行举手状态的识别。

本公开实施例中，第一得分存在一个数值区间，该区间内的数值的大小可以代表人体执行抬手动作的概率。

比如，设定第一得分 A 所属的区间为 $[0, 1]$ ，则第一得分 $A \in [0, 1]$ 。一般情况下，设定当 $A=0.23$ 时，即人体执行抬手动作的概率为 0.23% ，可以理解为，上述取值的第一得分对应的人体，
15 执行了属于非抬手动作的其他动作。再比如，设定当 $A=0.98$ 时，即人体执行抬手动作的概率为 0.98 ，可以理解为，上述取值的第一得分对应的人体，执行抬手的动作。

同样的，第二得分存在一个数值区间，该区间内的数值大小可以代表人体手部张开的概率。

比如，设定第二得分 B 所属的区间为 $[0, 1]$ ，则第二得分 $B \in [0, 1]$ 。一般情况下，设定当 $B=0.17$ 时，即为人体手部处于张开状态的概率为 0.17 ，可以理解为，上述取值的第二得分对应
20 的人体的手部处于非张开状态的其他状态。再比如，设定当 $B=0.93$ 时，即为人体手部处于张开状态的概率为 0.93 ，可以理解为，上述取值的第二得分对应的人体的手部处于张开状态。

进一步地，基于第一得分判断人体的上肢是否执行抬手动作，基于第二得分判断人体的手部是否处于张开状态，当确定人体的执行抬手动作的同时手部处于张开状态，则可以确定人体执行了举手动作。

25 本公开提出的举手识别方法，通过人体上半身的关键点的提取并生成第一图像特征，将第一图像特征输入人体动作识别模型进行抬手动作是否执行的检测，获取相应地第一得分。通过对上半身关键点进行筛选，获取人体手部的关键点并生成第二图像特征，将第二图像特征输入人体手部状态识别模型进行手部状态是否张开的检测，获取相应地第二得分。进一步地，基于第一得分与第二得分，实现对于人体举手动作的识别。本公开中，通过对于人体上半身关键点的提取，实现对于抬手动作的检测，通过对于手部关键点的采集，实现对于人体手掌是否张开的状态检测，
30 进一步地，融合对人体姿态以及人手状态的双重识别，使得可以在相对模糊且人体较小的远景图像中实现人体的举手动作识别，有效提高了人体举手识别的准确率。

上述实施例中，第一图像特征的获取方法可结合图3进一步理解，图3为本公开另一实施例的举手识别方法的流程示意图，如图3所示，该方法包括：

35 S301，将人体图像输入人体检测模型中进行人体检测，输出人体检测框。

为了实现对于第一图像特征的提取，需要设置一个人体检测模型，该模型可以基于采集到的待识别的人体图像生成人体检测框。

可选地，人体检测模型可以是轻量检测模型，其中，模型的主干网络（backbone）可以为卷积神经网络模型（ShuffleNet V2），检测框架可以为对象检测算法（You Only Look Once V3，YOLO V3）。依托于图像采集设备获取到的待识别的人体图像，将其输入至设定的人体检测模型后，可以生成人体检测框（Body_boxex）。
40

S302，将人体检测框输入人体关键点检测模型中进行特征提取，基于提取的特征进行关键点检测，输出上半身关键点。

为了实现人体的关键点的提取，需要将人体检测框输入人体关键点检测模型。本公开实施例中，基于人体检测框对应的人体组织结构部位的不同，分别对人体不同的部位进行关键点提取。
45

人体检测框基于待识别图像中的人体图像生成，包含有人体上半身的对应的不同部位的待采

集关键点的样本图像，将人体检测框输入至人体关键点检测模型后，可以获取到人体上半身不同部位对应的关键点。

进一步地，将每个部位对应的关键点组合，获取到所需提取的全部的人体上半身的关键点。

S303，基于上半身关键点的位置信息和置信度，构建输入矩阵。

5 一般情况下，统计学中针对某一个概率样本存在对应的置信区间，该置信区间可以理解为预设的概率样本可能的取值范围，进一步地，该概率样本预设的取值范围的准确程度，可以理解为置信度。

为了保证提取到的人体上半身关键点的所处位置的准确性，获取全部上半身的关键点后，需要对关键点的位置信息进行进一步的确认，可选地，可以通过人体上半身关键点的位置信息的置信度实现。

10 本公开实施例中，提取到的上半身关键点分别对应不同的人体上半身的组织结构，根据人体的基础结构，可以获取每个组织结构对应的关键点的正确位置信息，基于该位置信息，可以获取每个关键点的位置信息对应的置信区间。

进一步地，基于实际提取的人体上半身的关键点的位置信息与获取到的每个关键点对应的置信区间，确定每个人体上半身关键点的置信度。

可以为每个关键点设置一个对应的置信度阈值，当提取到的上半身的各个关键点均大于其对应的置信度阈值时，可以确定当前提取到的人体上半身的关键点位置准确，为有效可用的关键点。

进一步地，基于采集到的人体上半身的关键点的位置信息以及对应的置信度，构建输入矩阵。其中，构建矩阵的格式根据预设的人体动作识别模型的输入格式设置。

20 可选地，可以使用图卷积神经网络模型（Graph Convolutional Network，GCN）实现人体动作的识别。根据 GCN 分类模型输入矩阵的要求，可以构建一个 16×3 的输入矩阵，其中，16 代表每个输入矩阵中所包含的关键点数据的数量为 16 个，3 代表输入矩阵中每个关键点数据对应的坐标信息（x，y）的两个数值以及每个关键点对应的置信度。

S304，基于上半身关键点之间的连接关系，构建邻接矩阵和度矩阵。

25 确定上半身关键点的位置信息后，可以获取关键点与关键点之间的连接关系。一般情况下，上半身关键点存在一个或者多个可以与其相连接的边。

本公开实施例中，第一图像特征基于提取到的关键点的位置信息以及关键点与关键点之间的连接关系生成，因此，需要准确的提取表现关键点与关键点之间的连接关系。

进一步地，上半身的关键点与关键点之间的连接关系，可以用邻接矩阵体现。邻接矩阵的逻辑结构可以分为两个部分，其中一部分是顶点，在本公开实施例中即为提取到的上半身的关键点对应的数据，可以用一个一维数组存放。另一部分是边，在本公开实施例中即为可以体现上半身的关键点与关键点之间的连接关系数据，可以用一个二维数组存放。其中，存放上半身的关键点与关键点之间连接关系的二维数组，即为本公开实施例需要构建的邻接矩阵。

35 一般情况下，当两个关键点之间存在直接连接关系时，则该连接关系对应邻接矩阵中的元素的数值为 1，当两个关键点之间不存在直接连接的关系时，则该连接关系对应邻接矩阵中的元素的数值为 0。基于上述设定，可以获取全部的上半身关键点与关键点之间的连接关系对应的邻接矩阵。

进一步地，基于确定的上半身关键点与关键点之间的连接关系，可以构建度矩阵，度矩阵可以用于体现与提取到的关键点存在连接关系的边的数量。一般情况下，度矩阵是对角矩阵，其中所包括的顶点即为本公开实施例中的上半身的关键点的数据，对角上的各个元素代表了各个顶点的度，即为本公开实施例中的与各个上半身关键点存在连接关系的边的数量。基于上述设定，可以获取与全部的上半身关键点存在连接关系的边的数量对应的度矩阵。

S305，基于输入矩阵、度矩阵和邻接矩阵，获取第一图像特征。

45 本公开实施例中，输入矩阵可以确定提取到的上半身各个关键点的位置信息，邻接矩阵可以确定上半身各个关键点之间的连接关系，度矩阵可以确定与上半身每个关键点存在连接关系的边的数量。综合上述三个矩阵可以生成上半身关键点对应的图像特征，为第一图像特征。

进一步地，第一图像特征可以基于如下公式获取：

$$F_{\text{img}} = D^{-1}AX$$

其中，X 为输入矩阵、A 为邻接矩阵、D 为度矩阵，邻接矩阵与度矩阵所表征的上半身关键点与关键点之间的连接关系，对应人体的关节与关节之间的骨骼连接关系。

本公开提出的举手识别方法，通过构建输入矩阵、邻接矩阵与度矩阵，准确表现了提取到的上半身关键点的所处位置以及相互之间的连接关系，进一步地，基于输入矩阵、邻接矩阵与度矩阵生成第一图像特征，提高了第一图像特征的准确程度，进而使得基于第一图像特征获取的第一得分的准确程度得以保证，使得人体的抬手动作可以被精准识别。

上述实施例中，关于第二图像特征的提取可结合图 4 进一步理解，图 4 为本公开另一实施例的举手识别方法的流程示意图，如图 4 所示，该方法包括：

S401，基于关键点的位置信息，从手部关键点中确定出手肘关键点和手腕关键点。

本公开实施例中，人体的手部关键点可以包括人体的手肘对应的关键点、手腕对应的关键点以及手指对应的关键点等等。获取到全部的人体上半身的关键点后，可以基于关键点与关键点的位置信息，从全部的上半身关键点中获取人体手部的关键点。

基于每个手部关键点对应的的手部组织结构，确定各个手部关键点的位置信息，进一步地，基于关键点与关键点之间的相对位置信息，确定手肘关键点，以及手腕关键点。

S402，获取手肘关键点和手腕关键点的置信度，确定手肘关键点和手腕关键点的置信度均大于设定阈值。

获取到手肘关键点以及手腕关键点后，需要对其位置信息进行进一步的确定。根据人体的基础结构，可以获取手肘关键点与手腕关键点的正确位置信息，可以基于该位置信息获取手肘关键点与手腕关键点的对应的位置信息的置信区间。进一步地，基于实际获取到的手肘关键点与手腕关键点的位置信息，获取手肘关键点与手腕关键点的置信度。

进一步地，可以设置一个对应的置信度阈值，当手肘关键点与手腕关键点的置信度大于设定阈值时，可以理解为，手肘关键点与手腕关键点的所处位置是合理正确的。

进一步地，确定获取置信度均大于预设的置信度阈值的手肘关键点以及手腕关键点。

S403，基于手肘关键点和手腕关键点确定手部检测框，并从手部检测框内提取第二图像特征。

基于手肘关键点以及手腕关键点，可以进一步确定人体手掌的位置，进而生成人体的手部检测框。可以将手部检测框输入人体手部状态检测模型，进而识别人体的手掌是否处于张开状态。

本公开实施例中，手肘关键点与手腕关键点之间的连接关系，可以对应人体肘关节与腕关节之间的骨骼连接关系。

一般情况下，每个人的手掌长度与前臂长度存在一定的比例关系，同样地，人体的手肘关键点、手腕关键点之间的连接关系与手腕关键点、最长的手指终点对应的关键点之间的连接关系存在比例关系，二者相同。

获取到手肘关键点以及手腕关键点之间的连接关系后，可以基于上述比例关系，获取到人体的手腕关键点与最长的手指终点对应的关键点之间的连接关系，基于上述两个连接关系可以生成人体的手部检测框。

进一步地，基于手部检测框内包含的人体手部的关键点，生成人体手部的第二图像特征。

其中，第二图像特征可以用于识别人体手部是否处于张开状态，以及处于张开状态的程度。

本公开提出的举手识别方法，通过手肘关键点以及手腕关键点，确定人体的手部检测框，进而从手部检测框中获取人体手部的第二图像特征。进一步地，可以基于第二图像特征实现人体手部是否处于张开状态的识别，从而有效提高举手识别的准确程度。

上述实施例中，手部检测框的获取方法可结合图 5 进一步理解，图 5 为本公开另一实施例的举手识别方法的流程示意图，如图 5 所示，该方法包括：

S501，生成手肘关键点与手腕关键点之间的连线。

确定手肘关键点与手腕关键点以后，将二者之间相连接，进而获取手肘关键点与手腕关键点之间的连接关系，其中，该连接关系对应人体的肘关节与腕关节之间的骨骼连接关系，为直线。

S502, 沿着从手肘关键点向手腕关键点方向延伸连线设定距离, 以确定手指终点。

根据人体的基础结构, 人体的肘关节与腕关节之间以前臂相连接, 从手肘往手腕方向, 基于前臂的延长线上, 从手腕延伸出一定的距离, 即为人体的手掌部分。一般情况下, 人体的前臂长度与手掌长度存在比例关系。

- 5 相应地, 确定手肘关键点与手腕关键点之间的直线连接关系后, 沿手肘关键点至手腕关键点方向, 基于两个关键点之间的直线连接关系进行延长, 并延长至设定距离, 该设定距离为人体手腕关键点至最长的手指终点对应的关键点之间的距离, 与人体手掌长度相同。

可以通过手肘关键点与手腕关键点之间的直线连接关系的长度确定设定距离的值。

- 10 进一步地, 根据手肘关键点与手腕关键点的位置, 获取手肘关键点与手腕关键点之间的距离, 并基于上述前臂长度与手掌长度的比例关系, 确定设定距离。

比如, 设定手肘关键点为 (x_1, y_1) , 手腕关键点为 (x_2, y_2) , 则将 (x_1, y_1) 与 (x_2, y_2) 之间用直线相连接, 再基于手肘关键点至手腕关键点方向进行延长。设定手肘至手腕之间的连接线长度为 d_1 , 手腕至指尖的连接线中最长的长度为 d_2 , 则 d_1 与 d_2 之间的比例关系 $p = \frac{d_1}{d_2}$, 设定 p

值为 $\frac{1}{2}$ 。

- 15 进一步地, 基于如下的两点公式, 可以获取到人体手部的最长手指的手指终点 (x, y) :

$$\frac{x - x_2}{x_2 - x_1} = \frac{y - y_2}{y_2 - y_1}$$

S503, 基于手腕关键点和手指终点, 生成手部检测框。

一般情况下, 人体手部结构的手指终点可以多个, 不同手指的终点与手腕关键点之间的距离是不同的。

- 20 本公开实施例中, 手部检测框可以是具有规则形状的检测框, 比如矩形等等。获取到手肘关键点、手腕关键点以及最长的手指终点对应的关键点后, 可以确定全部的手部关键点的位置信息, 基于所有的手部关键点的位置信息, 生成手部检测框。

- 比如, 设定手部检测框为矩形, 则将手肘关键点至最长的手指终点对应的关键点之间的距离作为矩形的长, 将手掌张开状态下的拇指终点对应的关键点与小指终点对应的关键点之间的距离作为矩形的宽, 生成手部检测框。

本公开提出的举手识别方法, 通过手肘关键点与手腕关键点之间的延长线获取手指终点, 进一步基于手腕关键点与手指终点生成手部检测框, 使得可以基于手部检测框内的手部关键点生成人体手部的第二图像特征, 进而实现对于人体手部是否处于张开状态的检测识别, 有效提高了举手识别的准确程度。

- 30 本公开实施例提出的举手识别方法, 融合了人体抬手动作的识别以及人体手掌张开状态的识别。其中, 基于人体上半身关键点提取到的可以用于抬手动作检测的第一图像特征获取第一得分, 基于人体手部关键点提取到的可以用于手掌张开状态检测的第二图像特征获取第二得分, 进一步地, 基于第一得分和第二得分可以进行准确的举手识别。

- 可选地, 可以分别对第一得分和第二得分进行分析, 以获取检测结果, 如图 6 所示, 图 6 为本公开另一实施例的举手识别方法的流程示意图, 该方法包括:

S601, 响应于第一得分和第二得分均大于各自的第一预设得分阈值, 则确定人体处于举手状态。

第一得分与第二得分分别存在一个对应的得分阈值, 为第一预设得分阈值。

- 40 当第一得分大于其对应的第一预设得分阈值时, 可以理解为, 当前的第一图像特征中, 人体可以被确定执行了抬手动作。若第一得分小于或者等于其对应的第一预设得分阈值, 可以理解为,

当前状态下人体执行抬手动作的概率较小，且执行了其他非抬手的动作。

比如，设定第一得分与第一图像特征所属的采集图像中的人体手臂抬起后与桌面之间的角度为对应关系，第一得分所属区间为[0, 1]。

设定第一得分对应的第一预设得分阈值为 0.5，则该分数与人体手臂抬起后与桌面之间角度为 45°的状态相对应，可以理解为，当人体手臂抬起后与桌面之间角度大于 45°时，第一得分大于 0.5，进而可以确认人体执行了抬手动作。进一步地，设定当前第一得分为 0.8，则第一得分大于其对应的第一得分阈值 0.5，也就是说，当前状态下人体手臂抬起后与桌面之间的角度大于 45°，根据对应关系可知，第一得分为 0.8 时人体手臂抬起与桌面之间的角度为 72°，可以确定，当前状态下，第一图像特征所属的采集图像中，人体执行了抬手动作。

同样地，当第二得分大于其对应的第一预设阈值时，可以理解为，当前状态下人体的手掌确定处于张开的状态。若第二得分小于或者等于其对应的第二得分阈值，可以理解为，当前状态下人体的手掌处于非张开状态的其他状态。

比如，设定第二得分所属区间为[0, 1]，第二得分对应的第一预设得分阈值为 0.5，可以理解为，当第二得分大于 0.5 时，人体的手掌确定处于张开状态。

进一步地，当第一得分大于其对应的第一预设得分阈值，同时第二得分大于其对应的第一预设得分阈值时，可以判断，当前的人体手臂执行了抬手动作且手掌处于张开状态，可以确定该状态下的人体执行了举手动作。

本公开提出的举手识别方法，通过对第一得分的分析，确定人体是否执行了抬手动作，通过对第二得分的分析，确定人体手部是否处于张开状态，进而将确定执行抬手动作且手部处于张开状态的动作确定识别为举手动作，融合了对人体姿态以及人手状态的双重识别，使得可以在相对模糊且人体较小的远景图像中实现人体的举手动作识别，有效提高了人体举手识别的准确率。

可选地，可以融合第一得分与第二得分一起分析，以获取检测结果，如图 7 所示，图 7 为本公开另一实施例的举手识别方法的流程示意图，该方法包括：

S701，对第一得分和第二得分进行融合，生成第三得分，响应于第三得分大于第二预设得分阈值，则确定人体处于举手状态。

将第一得分与第二得分进行融合，生成第三得分。并设置一个第三得分对应的第二预设得分阈值，当第三得分大于第二预设得分阈值时，可以确定，当前状态下人体执行了抬手动作同时手掌处于张开状态，进而可以确定该状态下人体执行了举手动作。

可选地，可以为第一得分与第二得分配置权重，基于权重值进行第一得分与第二得分的融合，以生成第三得分。其中，第三得分可以体现人体执行抬手动作的同时手掌处于张开状态的概率。

比如，设定第三得分的所属区间为[0, 1]。设定第一得分为 0.8，权重占比为 0.7，第二得分为 0.7，权重占比为 0.3，则第三得分可以为 $0.8 \times 0.7 + 0.7 \times 0.3 = 0.56 + 0.21 = 0.77$ ，可以理解为，第三得分为 0.77 时所对应的人体执行抬手动作同时手掌处于张开状态的概率为 0.77。进一步地，设定第二预设得分阈值为 0.6，则当第三得分大于 0.6 时，可以确定人体执行了抬手动作同时手掌处于张开状态，当第三得分小于或者等于 0.6 时，可以确定人体并未执行举手动作。本示例中，获取到第三得分为 0.77，则第三得分大于第二预设得分阈值 0.6，可以确定，当前状态下人体执行了举手动作。

本公开提出的举手识别方法，基于第一得分与第二得分获取第三得分，通过对第三得分的分析，确定执行了抬手动作的同时手掌处于张开状态的动作，从而实现对于举手动作的识别，融合了人体的抬手动作的检测以及人体手掌部位张开状态的检测，有效提高了举手识别的准确程度。

与上述几种实施例提供的举手识别方法相对应，本公开的一个实施例还提供一种举手识别装置，由于本公开实施例提供的举手识别装置与上述几种实施例提供的举手识别方法相对应，因此在举手识别方法的实施方式也适用于本公开实施例提供的举手识别装置，在下述实施例中不再详细描述。

图 8 为本公开一实施例的举手识别装置的结构示意图，如图 8 所示，举手识别装置 800，包括提取模块 81、识别模块 82，其中：

提取模块 81，用于从采集的人体图像中提取人体的上半身关键点，并基于上半身关键点提取第一特征图像，输入人体动作识别模型中，输出被识别为人体执行抬手动作的第一得分；

提取模块 81，还用于从上半身关键点中确定手部关键点；

提取模块 81，还用于基于手部关键点，提取人手的第二图像特征，输入人手分类模型中，
5 输出被识别为手掌处于张开状态的第二得分；

识别模块 82，用于根据第一得分和第二得分进行举手状态的识别。

图 9 为本公开另一实施例的举手识别装置的结构示意图，如图 9 所示，举手识别装置 900，包括提取模块 91、识别模块 92，其中：

需要说明的是，提取模块 81、识别模块 82 与提取模块 91、识别模块 92，具备相同的结构
10 和功能。

本公开实施例中，提取模块 91，还用于：基于关键点的位置信息，从手部关键点中确定出手肘关键点和手腕关键点；基于手肘关键点和手腕关键点确定手部检测框，并从手部检测框内提取第二图像特征。

本公开实施例中，提取模块 91，还包括：

15 连接单元 911，用于生成手肘关键点与手腕关键点之间的连线；

延伸单元 912，用于沿着从手肘关键点向手腕关键点的方向延伸连线设定距离，以确定手指
20 终点；

生成单元 913，用于基于手腕关键点和手指终点，生成手部检测框。

本公开实施例中，提取模块 91，还用于：根据手肘关键点与手腕关键点的位置，获取手肘
25 关键点与手腕关键点之间的距离，根据距离确定设定距离。

本公开实施例中，提取模块 91，还用于：获取手肘关键点和手腕关键点的置信度，确定手肘关键点和手腕关键点的置信度均大于设定阈值。

本公开实施例中，提取模块 91，还用于：将人体图像输入人体检测模型中进行人体检测，
30 输出人体检测框；将人体检测框输入人体关键点检测模型中进行特征提取，基于提取的特征进行
关键点检测，输出上半身关键点。

本公开实施例中，提取模块 91，还包括：

矩阵构建单元 914，用于基于上半身关键点的位置信息和置信度，构建输入矩阵；

矩阵构建单元 914，还用于基于上半身关键点之间的连接关系，构建邻接矩阵和度矩阵；

特征提取单元 915，用于基于输入矩阵、度矩阵和邻接矩阵，获取第一图像特征。

35 本公开实施例中，识别模块 92，还用于：响应于第一得分和第二得分均大于各自的第一预设得分阈值，则确定人体处于举手状态；或者，对第一得分和第二得分进行融合，生成第三得分，
响应于第三得分大于第二预设得分阈值，则确定人体处于举手状态。

本公开提出的举手识别装置，提取采集图像中的人体的上半身关键点，生成第一图像特征，
40 将第一图像特征输入人体动作识别模型进行抬手动作检测，获取第一得分。通过人体的手部关键
点提取第二图像特征，将第二图像特征输入人体手部动作识别模型进行手掌张开状态检测，获取
第二得分。进一步地，基于第一得分与第二得分，实现对于人体是否执行举手动作的识别。本公
开中，融合了对人体姿态以及人手状态的双重识别，使得可以在相对模糊且人体较小的远景图像
中实现人体的举手动作识别，从而有效提高人体举手识别的准确率。

图 10 示出了可以用来实施本公开的实施例的示例电子设备 1000 的示意性框图。电子设备旨
45 在表示各种形式的数字计算机，诸如，膝上型计算机、台式计算机、工作台、个人数字助理、服
务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动
装置，诸如，个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所
示的部件、它们的连接和关系、以及它们的功能仅作为示例，并且不意在限制本文中描述的和/
或者要求的本公开的实现。

如图 10 所示，设备 1000 包括计算单元 1001，其可以根据存储在只读存储器（ROM）1002
中的计算机程序或者从存储单元 1008 加载到随机访问存储器（RAM）1004 中的计算机程序，来

执行各种适当的动作和处理。在 RAM 1004 中，还可存储设备 1000 操作所需的各种程序和数据。计算单元 1001、ROM 1002 以及 RAM 1004 通过总线 1004 彼此相连。输入/输出 (I/O) 接口 1005 也连接至总线 1004。

设备 1000 中的多个部件连接至 I/O 接口 1005，包括：输入单元 1006，例如键盘、鼠标等；
5 输出单元 1007，例如各种类型的显示器、扬声器等；存储单元 1008，例如磁盘、光盘等；以及通信单元 1009，例如网卡、调制解调器、无线通信收发机等。通信单元 1009 允许设备 1000 通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

计算单元 1001 可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元 1001 的一些示例包括但不限于中央处理单元 (CPU)、图形处理单元 (GPU)、各种专用的人工智能 (AI) 计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器 (DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元 1001 执行上文所描述的各个方法和处理，例如举手识别方法。例如，在一些实施例中，举手识别方法可被实现为计算机软件程序，其被有形地包含于机器可读介质，例如存储单元 1008。在一些实施例中，计算机程序的部分或者全部可以经由 ROM 1002 和/或通信单元 1009 而被载入和/或安装到设备 1000 上。当计算机程序加载到 RAM 1004 并由
10 计算单元 1001 执行时，可以执行上文描述的举手识别方法一个或多个步骤。备选地，在其他实施例中，计算单元 1001 可以通过其他任何适当的方式 (例如，借助于固件) 而被配置为执行举手识别方法。

本文中以上描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列 (FPGA)、专用集成电路 (ASIC)、专用标准产品 (ASSP)、芯片上系统的系统 (SOC)、负载可编程逻辑设备 (CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。
20 这些各种实施方式可以包括：实施在一个或者多个计算机程序中，该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释，该可编程处理器可以是专用或者通用可编程处理器，可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令，并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器，使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行，作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

在本公开的上下文中，机器可读介质可以是有形的介质，其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备，或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器 (RAM)、只读存储器 (ROM)、可擦除可编程只读存储器 (EPROM 或快闪存储器)、光纤、便捷式紧凑盘只读存储器 (CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。
30 35

为了提供与用户的交互，可以在计算机上实施此处描述的系统和技术，该计算机具有：用于向用户显示信息的显示装置 (例如，CRT (阴极射线管) 或者 LCD (液晶显示器) 监视器)；以及键盘和指向装置 (例如，鼠标或者轨迹球)，用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互；例如，提供给用户的反馈可以是任何形式的传感反馈 (例如，视觉反馈、听觉反馈、或者触觉反馈)；并且可以用任何形式 (包括声输入、语音输入或者、触觉输入) 来接收来自用户的输入。

可以将此处描述的系统和技术实施在包括后台部件的计算系统 (例如，作为数据服务器)、或者包括中间件部件的计算系统 (例如，应用服务器)、或者包括前端部件的计算系统 (例如，具有图形用户界面或者网络浏览器的用户计算机，用户可以通过该图形用户界面或者该网络浏览
40 45

器来与此处描述的系统和技术的实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)、互联网和区块链网络。

- 5 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务端可以是云服务器,又称为云计算服务器或云主机,是云计算服务体系中的一项主机产品,以解决了传统物理主机与 VPS 服务(“Virtual Private Server”,或简称“VPS”)中,存在的管理难度大,业务扩展性弱的缺陷。服务端也可以为分布式系统的服务端,或者是结合区块链的服务器
- 10 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

- 15 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

权利要求书

1、一种举手识别方法，包括：

从采集的人体图像中提取人体的上半身关键点，并基于所述上半身关键点提取第一图像特征，输入人体动作识别模型中，输出被识别为人体执行抬手动作的第一得分；

从所述上半身关键点中确定手部关键点；

基于所述手部关键点，提取人手的第二图像特征，输入人手分类模型中，输出被识别为手掌处于张开状态的第二得分；

根据所述第一得分和所述第二得分进行举手状态的识别。

2、根据权利要求1所述的举手识别方法，其中，所述基于所述手部关键点，提取人手的第二图像特征，包括：

基于所述关键点的位置信息，从所述手部关键点中确定出手肘关键点和手腕关键点；

基于所述手肘关键点和手腕关键点确定手部检测框，并从所述手部检测框内提取所述第二图像特征。

3、根据权利要求2所述的举手识别方法，其中，所述基于所述手肘关键点和手腕关键点确定手部检测框，包括：

生成所述手肘关键点与所述手腕关键点之间的连线；

沿着从所述手肘关键点向所述手腕关键点的方向延伸所述连线设定距离，以确定手指终点；

基于所述手腕关键点和所述手指终点，生成所述手部检测框。

4、根据权利要求3所述的举手识别方法，其中，所述沿着从所述手肘关键点向所述手腕关键点的方向延伸所述连线设定距离之前，还包括：

根据所述手肘关键点与所述手腕关键点的位置，获取所述手肘关键点与所述手腕关键点之间的距离，根据所述距离确定所述设定距离。

5、根据权利要求2所述的举手识别方法，其中，所述基于所述手肘关键点和手腕关键点确定手部检测框之前，还包括：

获取所述手肘关键点和所述手腕关键点的置信度，确定所述手肘关键点和所述手腕关键点的置信度均大于设定阈值。

6、根据权利要求1所述的举手识别方法，其中，所述从采集的人体图像中提取人体的上半身关键点，还包括：

将所述人体图像输入人体检测模型中进行人体检测，输出人体检测框；

将所述人体检测框输入人体关键点检测模型中进行特征提取，基于提取的特征进行关键点检测，输出所述上半身关键点。

7、根据权利要求1或6所述的举手识别方法，其特征在于，所述基于所述上半身关键点提取第一图像特征，包括：

基于所述上半身关键点的位置信息和置信度，构建输入矩阵；

基于所述上半身关键点之间的连接关系，构建邻接矩阵和度矩阵；

基于所述输入矩阵、所述度矩阵和所述邻接矩阵，获取所述第一图像特征。

8、根据权利要求1-5任一项所述的举手识别方法，其中，所述根据所述第一得分和所述第二得分进行举手状态的识别，包括：

响应于所述第一得分和所述第二得分均大于各自的第一预设得分阈值，则确定人体处于举手状态；或者，

对所述第一得分和所述第二得分进行融合，生成第三得分，响应于所述第三得分大于第二预设得分阈值，则确定所述人体处于举手状态。

5

9、一种举手识别装置，包括：

提取模块，用于从采集的人体图像中提取人体的上半身关键点，并基于所述上半身关键点提取第一图像特征，输入人体动作识别模型中，输出被识别为人体执行抬手动作的第一得分；

所述提取模块，还用于从所述上半身关键点中确定手部关键点；

10 所述提取模块，还用于基于所述手部关键点，提取人手的第二图像特征，输入人手分类模型中，输出被识别为手掌处于张开状态的第二得分；

识别模块，用于根据所述第一得分和所述第二得分进行举手状态的识别。

10、根据权利要求9所述的举手识别装置，其中，所述提取模块，还用于：

15 基于所述关键点的位置信息，从所述手部关键点中确定出手肘关键点和手腕关键点；

基于所述手肘关键点和手腕关键点确定手部检测框，并从所述手部检测框内提取所述第二图像特征。

11、根据权利要求10所述的举手识别装置，其中，所述提取模块，还包括：

20 连接单元，用于生成所述手肘关键点与所述手腕关键点之间的连线；

延伸单元，用于沿着从所述手肘关键点向所述手腕关键点的方向延伸所述连线设定距离，以确定手指终点；

生成单元，用于基于所述手腕关键点和所述手指终点，生成所述手部检测框。

25 12、根据权利要求11所述的举手识别装置，其中，所述提取获取模块，还用于：

根据所述手肘关键点与所述手腕关键点的位置，获取所述手肘关键点与所述手腕关键点之间的距离，根据所述距离确定所述设定距离。

13、根据权利要求10所述的举手识别装置，其中，所述提取模块，还用于：

30 获取所述手肘关键点和所述手腕关键点的置信度，确定所述手肘关键点和所述手腕关键点的置信度均大于设定阈值。

14、根据权利要求9所述的举手识别装置，其中，所述提取模块，还用于：

将所述人体图像输入人体检测模型中进行人体检测，输出人体检测框；

35 将所述人体检测框输入人体关键点检测模型中进行特征提取，基于提取的特征进行关键点检测，输出所述上半身关键点。

15、根据权利要求9或14所述的举手识别装置，其特征在于，所述提取模块，还包括：

矩阵构建单元，用于基于所述上半身关键点的位置信息和置信度，构建输入矩阵；

40 所述矩阵构建单元，还用于基于所述上半身关键点之间的连接关系，构建邻接矩阵和度矩阵；

特征提取单元，用于基于所述输入矩阵、所述度矩阵和所述邻接矩阵，获取所述第一图像特征。

16、根据权利要求9-13任一项所述的举手识别装置，其中，所述识别模块，还用于：

响应于所述第一得分和所述第二得分均大于各自的第一预设得分阈值，则确定人体处于举手状态；

45 或者，

对所述第一得分和所述第二得分进行融合，生成第三得分，响应于所述第三得分大于第二预设得分

阈值，则确定所述人体处于举手状态。

17、一种电子设备，包括：

至少一个处理器；以及

5 与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行权利要求 1-8 中任一项所述的方法。

18、一种存储有计算机指令的非瞬时计算机可读存储介质，其中，所述计算机指令用于使所述计算

10 机执行根据权利要求 1-8 中任一项所述的方法。

19、一种计算机程序产品，包括计算机程序，所述计算机程序在被处理器执行时实现根据权利要求 1-8 中任一项所述的方法。

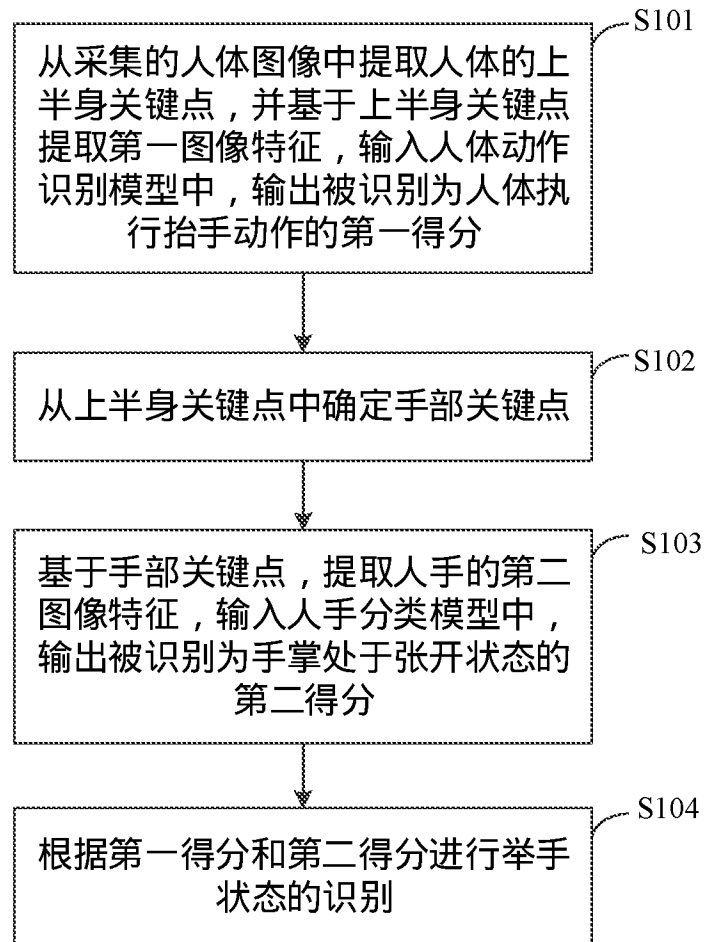


图 1

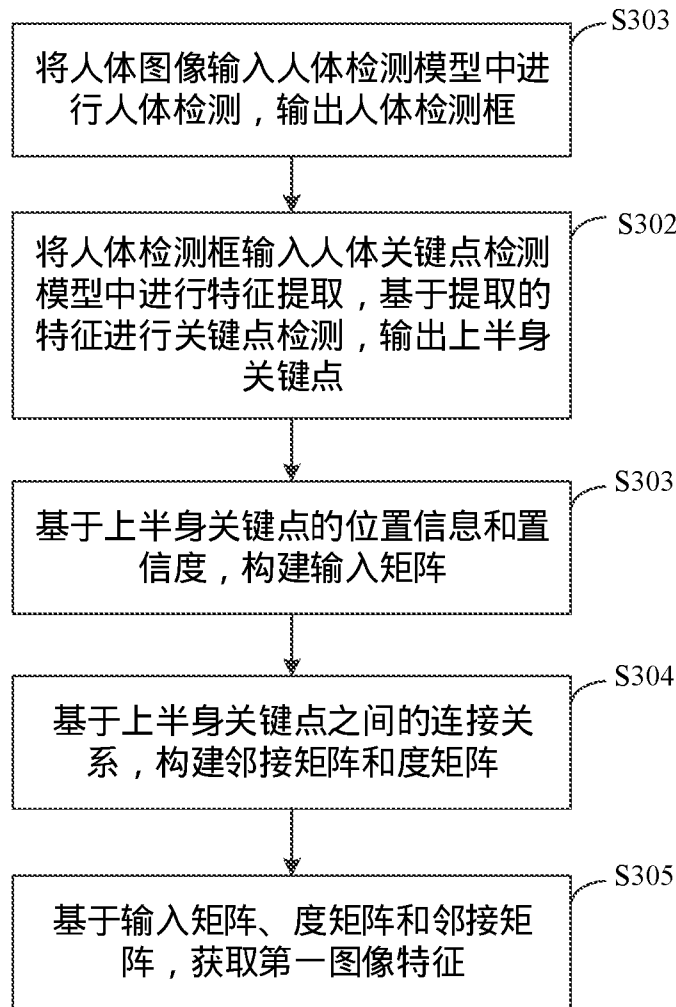


图 3

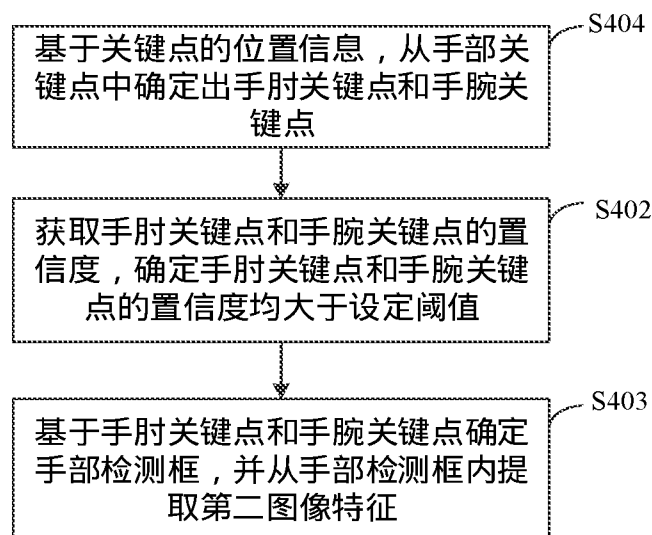


图 4

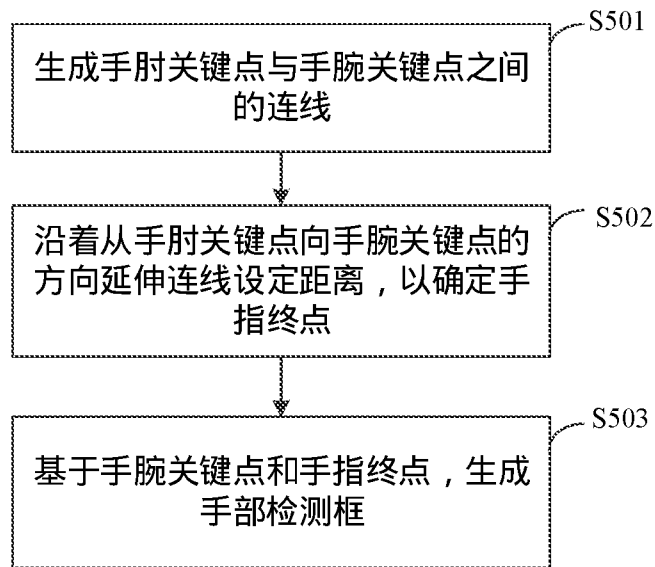


图 5

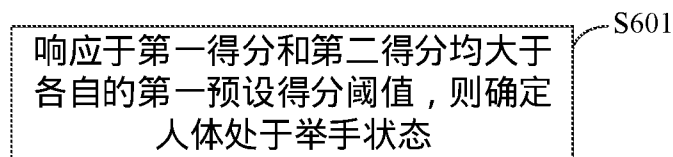


图 6

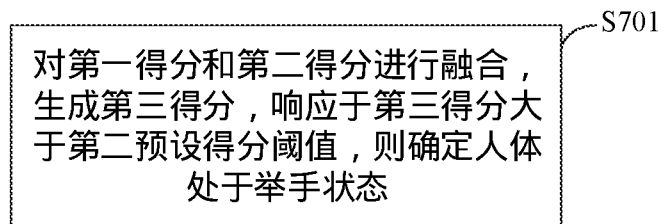


图 7

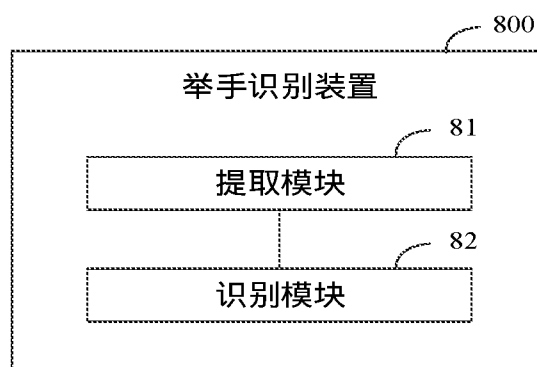


图 8

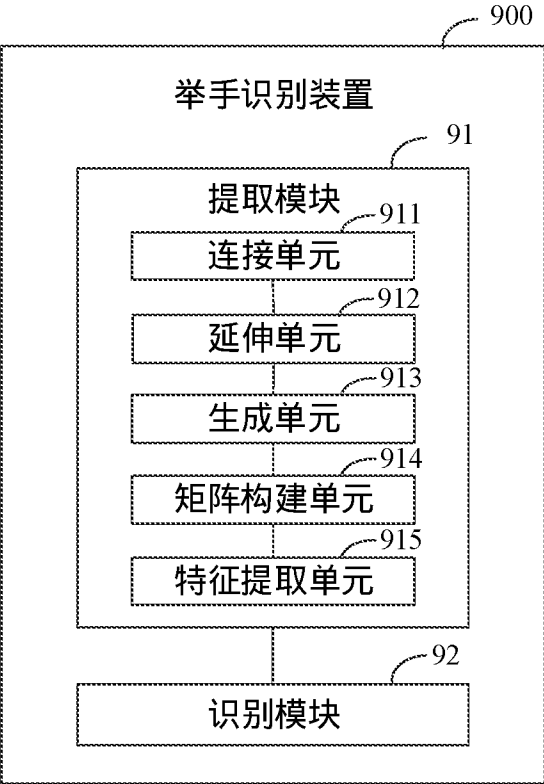


图 9

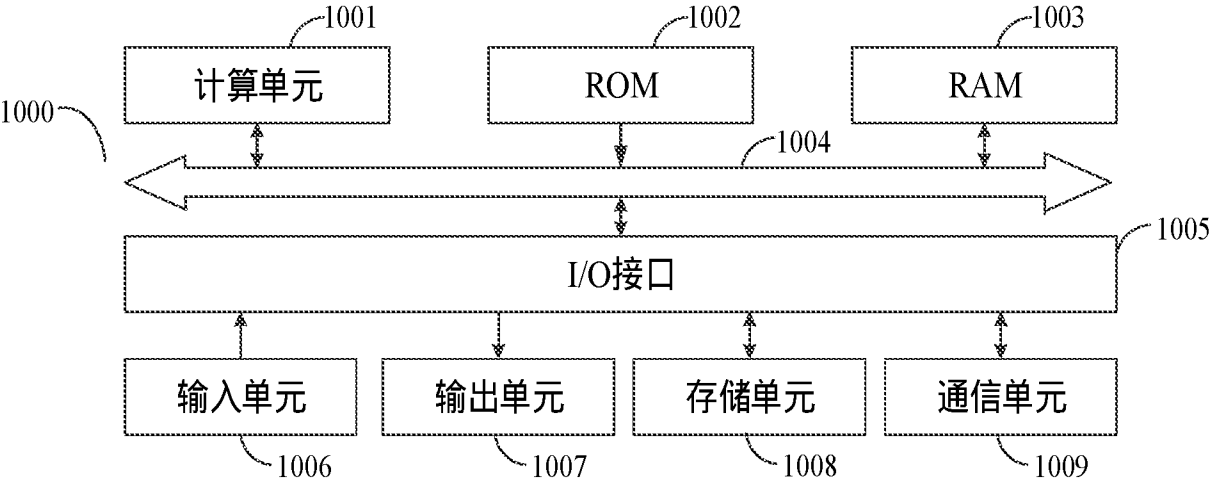


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/088032

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2022.01)i; G06K 9/62(2022.01)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)

G06K9/-; G06V40/-; G06N3/-

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)

CNABS; CNTXT; VEN; DWPI; ENTXT; USTXT; WOTXT; CNKI: 举手, 抬手, 识别, 检测, 上半身, 上身, 上肢, 关键点, 关节点, 特征点, 姿态, 位姿, 姿势, 手势, hands up, raise one's hand, recognize, identify, detect, upper body, upper limb, key point, articulation point, feature point, pose, posture, gesture

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113221745 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 06 August 2021 (2021-08-06) claims 1-19, and description, paragraphs 29-164	1-19
Y	CN 108038452 A (XIAMEN RUIWEI INFORMATION TECHNOLOGY CO., LTD.) 15 May 2018 (2018-05-15) description, paragraphs 14-30	1-19
Y	CN 111126339 A (BEIJING QIYI CENTURY SCIENCE & TECHNOLOGY CO., LTD.) 08 May 2020 (2020-05-08) description, paragraphs 32-143	1-19
Y	CN 112299172 A (CHENGDU XINCHAO MEDIA GROUP CO., LTD.) 02 February 2021 (2021-02-02) description, paragraphs 66-128	1-19
PY	CN 114360047 A (SHENZHEN HONGHE INNOVATION INFORMATION TECHNOLOGY CO., LTD.) 15 April 2022 (2022-04-15) description, paragraphs 70-152	1-19



Further documents are listed in the continuation of Box C.



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

20 June 2022

Date of mailing of the international search report

28 June 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/088032**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 110399822 A (SPIDER IOT TECHNOLOGY (BEIJING) CO., LTD.) 01 November 2019 (2019-11-01) description, paragraphs 25-121	1-19
A	US 2017315615 A1 (HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.) 02 November 2017 (2017-11-02) entire document	1-19

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/088032

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113221745	A	06 August 2021	None			
CN	108038452	A	15 May 2018	CN	108038452	B	03 November 2020
CN	111126339	A	08 May 2020	None			
CN	112299172	A	02 February 2021	None			
CN	114360047	A	15 April 2022	None			
CN	110399822	A	01 November 2019	None			
US	2017315615	A1	02 November 2017	WO	2016099561	A1	23 June 2016

国际检索报告

国际申请号

PCT/CN2022/088032

A. 主题的分类

G06K 9/00(2022.01)i; G06K 9/62(2022.01)i

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

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06K9/-; G06V40/-; G06N3/-

包含在检索领域中的除最低限度文献以外的检索文献

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))

CNABS;CNTXT;VEN;DWPI;ENTXT;USTXT;WOTXT;CNKI:举手, 抬手, 识别, 检测, 上半身, 上身, 上肢, 关键点, 关节点, 特征点, 姿态, 位姿, 姿势, 手势, hands up, raise one's hand, recognize, identify, detect, upper body, upper limb, key point, articulation point, feature point, pose, posture, gesture

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 113221745 A (北京百度网讯科技有限公司) 2021年8月6日 (2021 - 08 - 06) 权利要求1-19, 说明书第29-164段	1-19
Y	CN 108038452 A (厦门瑞为信息技术有限公司) 2018年5月15日 (2018 - 05 - 15) 说明书第14-30段	1-19
Y	CN 111126339 A (北京奇艺世纪科技有限公司) 2020年5月8日 (2020 - 05 - 08) 说明书第32-143段	1-19
Y	CN 112299172 A (成都新潮传媒集团有限公司) 2021年2月2日 (2021 - 02 - 02) 说明书第66-128段	1-19
PY	CN 114360047 A (深圳市鸿合创新信息技术有限责任公司) 2022年4月15日 (2022 - 04 - 15) 说明书第70-152段	1-19
Y	CN 110399822 A (思百达物联网科技北京有限公司) 2019年11月1日 (2019 - 11 - 01) 说明书第25-121段	1-19
A	US 2017315615 A1 (HEWLETT-PACKARD DEVELOPMENT COMPANY LP) 2017年11月2日 (2017 - 11 - 02) 全文	1-19

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“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)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“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

“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

“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

“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

“&” 同族专利的文件

国际检索实际完成的日期

2022年6月20日

国际检索报告邮寄日期

2022年6月28日

ISA/CN的名称和邮寄地址

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

传真号 (86-10)62019451

受权官员

罗湘

电话号码 86-(20)-28950390

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/088032

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	113221745	A	2021年8月6日	无	
CN	108038452	A	2018年5月15日	CN	108038452 B 2020年11月3日
CN	111126339	A	2020年5月8日	无	
CN	112299172	A	2021年2月2日	无	
CN	114360047	A	2022年4月15日	无	
CN	110399822	A	2019年11月1日	无	
US	2017315615	A1	2017年11月2日	WO	2016099561 A1 2016年6月23日