



(12) 发明专利

(10) 授权公告号 CN 114612764 B

(45) 授权公告日 2025. 02. 25

(21) 申请号 202210142734.8

G06V 20/60 (2022.01)

(22) 申请日 2022.02.16

G06V 10/74 (2022.01)

(65) 同一申请的已公布的文献号

G06V 10/774 (2022.01)

申请公布号 CN 114612764 A

G06V 10/82 (2022.01)

(43) 申请公布日 2022.06.10

G06N 3/045 (2023.01)

(73) 专利权人 厦门瑞为信息技术有限公司

G06N 3/0464 (2023.01)

地址 361000 福建省厦门市思明区软件园

G06N 3/084 (2023.01)

二期望海路37号4楼

(56) 对比文件

(72) 发明人 陈敏达 徐绍凯 张帅 王汉超

CN 112884085 A, 2021.06.01

贾宝芝 何一凡

WO 2020134848 A1, 2020.07.02

(74) 专利代理机构 厦门市首创君合专利事务所

有限公司 35204

专利代理师 连耀忠

(51) Int. Cl.

G06V 40/10 (2022.01)

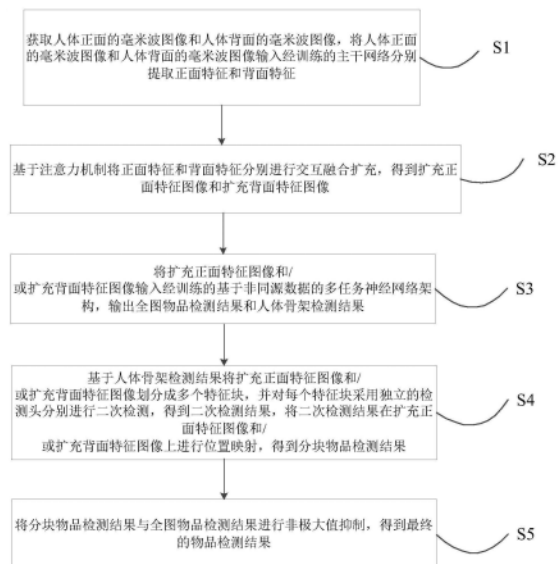
权利要求书3页 说明书10页 附图7页

(54) 发明名称

一种基于毫米波图像的物品检测方法、装置及可读介质

(57) 摘要

本发明公开了一种基于毫米波图像的物品检测方法、装置及可读介质,通过获取人体正面和背面的毫米波图像,将人体正面和背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征;基于注意力机制将正面特征和背面特征分别进行交互融合扩充,并将得到的扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果;基于人体骨架检测结果将图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行检测,得到二次检测结果,再在原图上进行位置映射,得到分块物品检测结果;将分块物品检测结果与全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。



1. 一种基于毫米波图像的物品检测方法,其特征在于,包括以下步骤:

S1,获取人体正面的毫米波图像和人体背面的毫米波图像,将所述人体正面的毫米波图像和人体背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征;

S2,基于注意力机制将所述正面特征和背面特征分别进行交互融合扩充,得到扩充正面特征图像和扩充背面特征图像,具体包括:

将由所述正面特征生成的查询向量与由背面特征生成的键向量进行相似度比对,再经softmax操作生成第一注意力矩阵,并将所述第一注意力矩阵对所述背面特征的值向量进行采样,将采样后的背面特征进行翻转并与所述正面特征进行拼接,得到扩充正面特征图像;

将由所述背面特征生成的查询向量与由正面特征生成的键向量进行相似度比对,再经softmax操作生成第二注意力矩阵,并将所述第二注意力矩阵对所述正面特征的值向量进行采样,将采样后的正面特征进行翻转并与所述背面特征进行拼接,得到扩充背面特征图像;

S3,将所述扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果;

S4,基于所述人体骨架检测结果将所述扩充正面特征图像和/或扩充背面特征图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行二次检测,得到二次检测结果,将所述二次检测结果在所述扩充正面特征图像和/或扩充背面特征图像上进行位置映射,得到分块物品检测结果;

S5,将所述分块物品检测结果与所述全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。

2. 根据权利要求1所述的基于毫米波图像的物品检测方法,其特征在于,所述主干网络采用swin-transformer结构。

3. 根据权利要求1所述的基于毫米波图像的物品检测方法,其特征在于,所述基于非同源数据的多任务神经网络架构包括分别对应于物品检测任务、人体骨架检测任务和其它视觉任务的任务头。

4. 根据权利要求3所述的基于毫米波图像的物品检测方法,其特征在于,所述步骤S3的基于非同源数据的多任务神经网络架构的训练过程具体包括:

通过采样器对非同源数据集进行采样,获取所述非同源数据集中每个样本的任务响应向量,所述任务响应向量对应于所述数据集中物品检测任务、人体骨架检测任务和其它视觉任务的标注情况,并将每个样本的所述任务响应向量进行拼接得到任务响应矩阵;

将所述非同源数据集对应的所述扩充正面特征图像和扩充背面特征图像输入具有多个任务头的基于非同源数据的多任务神经网络架构,经过多个所述任务头分别进行前向传播,生成每个任务头对应的损失值,将所有任务头对应的损失值拼接成损失向量;

将所述任务响应矩阵和损失向量点乘后得到最终的损失,并进行反向传播,直至达到预期效果或满足训练结束条件。

5. 根据权利要求1所述的基于毫米波图像的物品检测方法,其特征在于,所述步骤S4具体包括:

S41,根据所述人体骨架检测结果将所述扩充正面特征图像和/或扩充背面特征图像划

分成多个区域；

S42,基于每个区域的坐标通过RoIAlign操作对所述扩充正面特征图像和/或扩充背面特征图像进行采样,从而获得多个特征块,并记录每个特征块在所述扩充正面特征图像和/或扩充背面特征图像中的起始位置；

S43,对每个特征块采用独立的检测头进行二次检测,得到每个特征块的二次检测结果；

S44,将所述每个特征块的二次检测结果中的检测框坐标分别加上每个特征块在所述扩充正面特征图像和/或扩充背面特征图像中的起始位置,得到所述分块物品检测结果。

6.根据权利要求5所述的基于毫米波图像的物品检测方法,其特征在于,在所述步骤S43中还包括响应于第一特征块和第二特征块具有对称部位,将所述第二特征块进行翻转后输入所述第一特征块对应的检测头进行二次检测。

7.根据权利要求1-6中任一项所述的基于毫米波图像的物品检测方法,其特征在于,所述物品包括违禁品。

8.一种基于毫米波图像的物品检测装置,其特征在于,包括:

数据获取模块,被配置为获取人体正面的毫米波图像和人体背面的毫米波图像,将所述人体正面的毫米波图像和人体背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征；

特征扩充模块,被配置为基于注意力机制将所述正面特征和背面特征分别进行交互融合扩充,得到扩充正面特征图像和扩充背面特征图像,具体包括:

将由所述正面特征生成的查询向量与由背面特征生成的键向量进行相似度比对,再经softmax操作生成第一注意力矩阵,并将所述第一注意力矩阵对所述背面特征的值向量进行采样,将采样后的背面特征进行翻转并与所述正面特征进行拼接,得到扩充正面特征图像；

将由所述背面特征生成的查询向量与由正面特征生成的键向量进行相似度比对,再经softmax操作生成第二注意力矩阵,并将所述第二注意力矩阵对所述正面特征的值向量进行采样,将采样后的正面特征进行翻转并与所述背面特征进行拼接,得到扩充背面特征图像；

全面检测模块,被配置为将所述扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果；

分块检测模块,被配置为基于所述人体骨架检测结果将所述扩充正面特征图像和/或扩充背面特征图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行二次检测,得到,二次检测结果,将所述二次检测结果在所述扩充正面特征图像和/或扩充背面特征图像上进行位置映射,得到分块物品检测结果；

合并模块,被配置为将所述分块物品检测结果与所述全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。

9.一种电子设备,包括:

一个或多个处理器；

存储装置,用于存储一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-7中任一所述的方法。

10.一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1-7中任一所述的方法。

一种基于毫米波图像的物品检测方法、装置及可读介质

技术领域

[0001] 本发明计算机视觉及深度学习领域,具体涉及一种基于毫米波图像的物品检测方法、装置及可读介质。

背景技术

[0002] 毫米波人体成像设备具有对人体无害、穿透力强的特点,其发射功率不及手机电磁波辐射的千分之一,能准确识别人体携带物品,有效提高检查的客观性、准确性、针对性,降低安检员的劳动强度,提升安检效率。

[0003] 毫米波由于其能穿透衣物又不伤害人体的特性,在现今的智能安检领域中大放异彩。毫米波技术与下游的机器视觉算法结合可以大幅度提高安检流程的效率和可靠性,减少安检过程的人力投入。然而当下的智能安检系统,通常因为违禁品成像不明显、特征挖掘不充分、算法性能不高等问题,导致检测表现差强人意,无法承担安检流程的主要责任,只能作为人工安检的补充手段。

[0004] 现有技术中的毫米波违禁品检测存在以下问题:

[0005] 1. 单个角度的毫米波扫描中,违禁品的成像通常不够完整,提取出的特征通常也不够丰富,从而导致这部分违禁品的检出率较低。

[0006] 2. 现有的检测方案通常采用一个模型处理一个视觉任务的模式。但当应用场景需要多种任务并行时,多个模型的推理会使算法速度会严重下降,导致算法在一些需要实时的场景无法达到速度要求或者对硬件要求十分苛刻。

[0007] 3. 受制于毫米波成像技术的发展,部分违禁品的成像不够清晰(如密度较低的粉末爆炸物,体积较小的打火机等)。这些违禁品时常与某些背景特征混淆(如绑带、纽扣、肌肉线条、成像噪声等),从而造成误检或漏检。

[0008] 4. 现有的违禁品检测方法,通常使用卷积神经网络提取特征,这种方法可以很好地获取目标物品的纹理信息,但是缺少对目标局部上下文的语义理解。而事实上,在违禁品检测的场景,这种局部语义理解是很有必要的,因为大部分违禁品周围都会出现口袋、胶带/绑带、衣服褶皱等特征,结合这些特征再进行检测,可以显著的提高准确率。

发明内容

[0009] 针对上述提到的技术问题。本申请的实施例的目的在于提出了一种基于毫米波图像的物品检测方法、装置及可读介质,来解决以上背景技术部分提到的技术问题。

[0010] 第一方面,本申请的实施例提供了一种基于毫米波图像的物品检测方法,包括以下步骤:

[0011] S1,获取人体正面的毫米波图像和人体背面的毫米波图像,将人体正面的毫米波图像和人体背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征;

[0012] S2,基于注意力机制将正面特征和背面特征分别进行交互融合扩充,得到扩充正面特征图像和扩充背面特征图像;

[0013] S3,将扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果;

[0014] S4,基于人体骨架检测结果将扩充正面特征图像和/或扩充背面特征图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行二次检测,得到二次检测结果,将二次检测结果在扩充正面特征图像和/或扩充背面特征图像上进行位置映射,得到分块物品检测结果;

[0015] S5,将分块物品检测结果与全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。

[0016] 在具体的实施例中,主干网络采用swin-transformer结构。

[0017] 在具体的实施例中,步骤S2具体包括:

[0018] 将由正面特征生成的查询向量与由背面特征生成的键向量进行相似度比对,再经softmax操作生成第一注意力矩阵,并将第一注意力矩阵对背面特征的值向量进行采用,将采样后的背面特征进行翻转并与正面特征进行拼接,得到扩充正面特征图像;

[0019] 将由背面特征生成的查询向量与由正面特征生成的键向量进行相似度比对,再经softmax操作生成第二注意力矩阵,并将第二注意力矩阵对正面特征的值向量进行采样,将采样后的正面特征进行翻转并与背面特征进行拼接,得到扩充背面特征图像。

[0020] 在具体的实施例中,基于非同源数据的多任务神经网络架构包括分别对应于物品检测任务、人体骨架检测任务和其它视觉任务的任务头。

[0021] 在具体的实施例中,步骤S3的基于非同源数据的多任务神经网络架构的训练过程具体包括:

[0022] 通过采样器对非同源数据集进行采样,获取非同源数据集中每个样本的任务响应向量,任务响应向量对应于数据集中物品检测任务、人体骨架检测任务和其它视觉任务的标注情况,并将每个样本的任务响应向量进行拼接得到任务响应矩阵;

[0023] 将非同源数据集对应的扩充正面特征图像和扩充背面特征图像输入具有多个任务头的基于非同源数据的多任务神经网络架构,经过多个任务头分别进行前向传播,生成每个任务头对应的损失值,将所有任务头对应的损失值拼接成损失向量;

[0024] 将任务响应矩阵和损失向量点乘后得到最终的损失,并进行反向传播,直至达到预期效果或满足训练结束条件。

[0025] 在具体的实施例中,步骤S4具体包括:

[0026] S41,根据人体骨架检测结果将扩充正面特征图像和/或扩充背面特征图像划分成多个区域;

[0027] S42,基于每个区域的坐标通过RoIAlign操作对扩充正面特征图像和/或扩充背面特征图像进行采样,从而获得多个特征块,并记录每个特征块在扩充正面特征图像和/或扩充背面特征图像中的起始位置;

[0028] S43,对每个特征块采用独立的检测头进行二次检测,得到每个特征块的二次检测结果;

[0029] S44,将每个特征块的二次检测结果中的检测框坐标分别加上每个特征块在扩充正面特征图像和/或扩充背面特征图像中的起始位置,得到分块物品检测结果。

[0030] 在具体的实施例中,在步骤S43中还包括响应于第一特征块和第二特征块具有对

称部位,将第二特征块进行翻转后输入第一特征块对应的检测头进行二次检测。

[0031] 在具体的实施例中,物品包括违禁品。

[0032] 第二方面,本申请的实施例提供了一种基于毫米波图像的物品检测装置,包括:

[0033] 数据获取模块,被配置为获取人体正面的毫米波图像和人体背面的毫米波图像,将人体正面的毫米波图像和人体背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征;

[0034] 特征扩充模块,被配置为基于注意力机制将正面特征和背面特征分别进行交互融合扩充,得到扩充正面特征图像和扩充背面特征图像;

[0035] 全面检测模块,被配置为将扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果;

[0036] 分块检测模块,被配置为基于人体骨架检测结果将扩充正面特征图像和/或扩充背面特征图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行二次检测,得到,二次检测结果,将二次检测结果在扩充正面特征图像和/或扩充背面特征图像上进行位置映射,得到分块物品检测结果;

[0037] 合并模块,被配置为将分块物品检测结果与全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。

[0038] 第三方面,本申请的实施例提供了一种电子设备,包括一个或多个处理器;存储装置,用于存储一个或多个程序,当一个或多个程序被一个或多个处理器执行,使得一个或多个处理器实现如第一方面中任一实现方式描述的方法。

[0039] 第四方面,本申请的实施例提供了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如第一方面中任一实现方式描述的方法。

[0040] 相比于现有技术,本发明具有以下有益效果:

[0041] (1) 本发明采用经训练的主干网络分别提取正面特征和背面特征,因此能够提供更加全面的检测物特征,提高物品的检出率。

[0042] (2) 本发明采用基于非同源数据的多任务神经网络架构,该模型中每个任务都具有地理的任务头进行检测,并且每个任务头的输入都可以复用主干网络所提取的特征,有效降低模型的训练难度,在不提升标注成本的前提下,大幅提升了多个任务并行处理的速度,使算法满足实时要求。

[0043] (3) 本发明还对被检测目标进行人体骨架检测,以便于后续根据人体骨架检测结果分块,再对分块后得到的多个特征块分别再经过一个独立的检测头进行二次检测,对于每个检测头来说输入的背景特征相对固定,进而减少误检并提高检出率。

[0044] (4) 本发明的主干网络采用了swin-transformer结构,引入了卷积网络才有的聚焦局部特征和形成层次化特征金字塔的能力,使得主干网络结合了时序网络和卷积网络的优点,从而大大提升了检测表现。

[0045] (5) 本发明大大提高了违禁品检测任务的准确率和效率,使由智能安检系统主导的安检流程成为可能。

附图说明

[0046] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简要介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域的普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0047] 图1是本申请的一个实施例可以应用于其中的示例性装置架构图;

[0048] 图2为本发明的实施例的基于毫米波图像的物品检测方法的流程示意图;

[0049] 图3为本发明的实施例的基于毫米波图像的物品检测方法的违禁品检测的整体流程示意图;

[0050] 图4为本发明的实施例的基于毫米波图像的物品检测方法的步骤S2的流程示意图;

[0051] 图5为本发明的实施例的基于毫米波图像的物品检测方法的基于非同源数据的多任务神经网络架构的训练过程示意图;

[0052] 图6为本发明的实施例的基于毫米波图像的物品检测方法的全图物品检测结果和人体骨架检测结果;

[0053] 图7为本发明的实施例的基于毫米波图像的物品检测方法的根据人体骨架检测结果划分区域的结果;

[0054] 图8为本发明的实施例的基于毫米波图像的物品检测装置的示意图;

[0055] 图9是适于用来实现本申请实施例的电子设备的计算机装置的结构示意图。

具体实施方式

[0056] 为了使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明作进一步地详细描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,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它实施例,都属于本发明保护的范围。

[0057] 图1示出了可以应用本申请实施例的基于毫米波图像的物品检测方法或基于毫米波图像的物品检测装置的示例性装置架构100。

[0058] 如图1所示,装置架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0059] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种应用,例如数据处理类应用、文件处理类应用等。

[0060] 终端设备101、102、103可以是硬件,也可以是软件。当终端设备101、102、103为硬件时,可以是各种电子设备,包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。当终端设备101、102、103为软件时,可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块(例如用来提供分布式服务的软件或软件模块),也可以实现成单个软件或软件模块。在此不做具体限定。

[0061] 服务器105可以是提供各种服务的服务器,例如对终端设备101、102、103上传的文

件或数据进行处理的后台数据处理服务器。后台数据处理服务器可以对获取的文件或数据进行处理,生成处理结果。

[0062] 需要说明的是,本申请实施例所提供的基于毫米波图像的物品检测方法可以由服务器105执行,也可以由终端设备101、102、103执行,相应地,基于毫米波图像的物品检测装置可以设置于服务器105中,也可以设置于终端设备101、102、103中。

[0063] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的数目。在所处理的数据不需要从远程获取的情况下,上述装置架构可以不包括网络,而只需服务器或终端设备。

[0064] 图2示出了本申请的实施例提供的一种基于毫米波图像的物品检测方法,包括以下步骤:

[0065] S1,获取人体正面的毫米波图像和人体背面的毫米波图像,将人体正面的毫米波图像和人体背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征。

[0066] 具体地,本申请的实施例中通过采集来自人体正面和背面两个角度的毫米波图像,以避免在单个角度的毫米波扫描的物品成像不够完整,所提取的特征也不够丰富,导致物品的检出率较低等问题。参考图3,在本申请的实施例中物品检测在安检场景下以违禁品检测为例,在其他实施例中适用于其他物品的检测。在具体的实施例中,主干网络采用swin-transformer结构。swin-transformer结构除了transformer固有的捕获全局依赖、强语义理解的能力外,该网络还通过分块(patching)、滑动窗口(windowing)、窗口偏移(shift window)等操作,引入了卷积网络才有的聚焦局部特征和形成层次化特征金字塔的能力,使得主干网络结合了时序网络和卷积网络的优点,从而大大提升了检测表现。主干网络可以选择任何带有多层次下采样的卷积神经网络,此处选择swin-transformer结构主要是因为观察到其结果最好,而且不管是从理论上还是实践上,其捕获局部依赖的能力确实解决了在违禁品检测的场景的局部语义理解相关的问题,因为大部分违禁品周围都会出现口袋、胶带/绑带、衣服褶皱等特征,结合这些特征再进行检测,可以显著的提高准确率。

[0067] S2,基于注意力机制将正面特征和背面特征分别进行交互融合扩充,得到扩充正面特征图像和扩充背面特征图像。

[0068] 在具体的实施例中,步骤S2具体包括:

[0069] 将由正面特征生成的查询向量与由背面特征生成的键向量进行相似度比对,再经softmax操作生成第一注意力矩阵,并将第一注意力矩阵对背面特征的值向量进行采用,将采样后的背面特征进行翻转并与正面特征进行拼接,得到扩充正面特征图像;

[0070] 将由背面特征生成的查询向量与由正面特征生成的键向量进行相似度比对,再经softmax操作生成第二注意力矩阵,并将第二注意力矩阵对正面特征的值向量进行采样,将采样后的正面特征进行翻转并与背面特征进行拼接,得到扩充背面特征图像。

[0071] 具体地,参考图4,以正面检测为例,在进行正面检测时,通过注意力机制对背面特征也进行采样,采样过程具体为:将由正面特征生成的查询向量(query)与由背面特征生成的键向量(key)进行相似度比对,再经softmax操作生成第一注意力矩阵,并用该注意力矩阵对背面特征的值向量(value)进行采样。然后将采样后的背面特征水平翻转并与正面特征进行拼接,得到扩充正面特征图像,最后将扩充正面特征图像进行下游任务检测。同样背面检测时采用相同的原理得到扩充背面特征图像。

[0072] 因此,本方法先用一个预先训练好的主干网络分别提取正面特征和背面特征,在对正面图像进行检测时,使用背面特征对正面特征进行扩充,反之亦然。这样可以提供更全面的检测物特征,尤其是针对人体侧面检测时正背面各只有一半成像的违禁品,从而提高这部分违禁品的检出率。

[0073] S3,将扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果。

[0074] 具体地,如果依照传统的神经网络训练方法,需要保证输入的所有训练数据都拥有全部任务的标注。在安检场景中,以违禁品检测为例,所有的数据都要进行违禁品检测、人体骨架检测、其他视觉任务三种标注。显然,人体骨架检测的任务难度远小于违禁品检测,训练这两个任务需要的数据量也远小于违禁品检测,因此对全部训练数据进行完全标注是性价比极低的。因此本方法采用了多任务的神经网络架构,即每个任务拥有自己独立的任务头,但每个任务头的输入都会复用主干网络提取的特征,即共享主干网络。

[0075] 在具体的实施例中,本方法采用物品检测、人体骨架检测的多任务神经网络架构以及非同源数据集的混合训练方法,基于非同源数据的多任务神经网络架构包括分别对应于物品检测任务、人体骨架检测任务和其它视觉任务的多个任务头。相比于传统的一个模型处理一个任务,本方法采用一个模型处理多个任务的架构,每个任务都复用主干网络中提取的特征,大大减少了神经网络的参数量和推理时间。

[0076] 为了避免冗余标注,本方法训练时使用非同源数据集,即带有不同类型标注的数据会被混合采样作为训练输入。本方法提出了能混合使用非同源数据集训练多任务神经网络架构的方法。在不提升标注成本的前提下,训练基于非同源数据的多任务神经网络架构,大幅提升了多个任务并行处理的速度,满足实时要求。本方法在采样时额外标记非同源数据集中每个样本的任务响应向量,而在计算损失时用任务响应向量来保证每个样本只计算其在响应任务中的损失值。在具体的实施例中,假定三个视觉任务:违禁品检测、人体骨架检测和其他视觉任务。为每个数据集设置一个任务响应向量,向量长度为任务数量,例如3,任务响应向量内的值由0或1组成。若该数据集有对应任务的标注,则任务响应向量的对应位置记为1,否则为0。例如:A数据集仅标注了违禁品信息,则其任务响应向量为[1 0 0];B数据集标注了人体骨架信息和其他视觉任务信息,则其任务响应向量为[0 1 1]。

[0077] 在具体的实施例中,步骤S3的基于非同源数据的多任务神经网络架构的训练过程具体包括:

[0078] 通过采样器对非同源数据集进行采样,获取非同源数据集中每个样本的任务响应向量,任务响应向量对应于数据集中物品检测任务、人体骨架检测任务和其它视觉任务的标注情况,并将每个样本的任务响应向量进行拼接得到任务响应矩阵;

[0079] 将非同源数据集对应的扩充正面特征图像和扩充背面特征图像输入具有多个任务头的基于非同源数据的多任务神经网络架构,经过多个任务头分别进行前向传播,生成每个任务头对应的损失值,将所有任务头对应的损失值拼接成损失向量;

[0080] 将任务响应矩阵和损失向量点乘后得到最终的损失,并进行反向传播,直至达到预期效果或满足训练结束条件。

[0081] 具体地,参考图5,采样器从非同源数据集中的各个数据集进行随机采样,并记录每个样本所属数据集的任务响应向量,采集一定数量(batch size)的样本后输入主干网络

进行特征提取,提取后的正面特征或背面特征经过步骤S2的特征扩充后输入经训练的基于非同源数据的多任务神经网络架构,基于非同源数据的多任务神经网络架构中的3个任务头分别进行前向传播,每个任务头生成一个损失值并组成损失向量L,损失向量L是一个大小为batch size*3的二维向量;同时将每个样本的任务响应向量拼接,得到任务响应矩阵M,任务响应矩阵M也是一个二维的任务响应向量M,大小同样为batch size*3,将损失向量L和任务响应矩阵M点乘后得到最终的损失,并进行反向传播。由于采用了任务响应向量,因此可以每个任务头的输入都可以复用同一个主干网络提取的特征,有效减少整个模型的规模,降低模型训练难度。基于非同源数据的多任务神经网络架构指的是采用上述的训练方法所得到的多个任务头组成的神经网络架构,并非具体某个神经网络模型,具体的每个任务头可以根据实际检测需求选取合适的现有的检测神经网络模型。

[0082] 将实时采集到的人体正面的毫米波图像和人体背面的毫米波图像所提取的扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果,如图6所示。

[0083] S4,基于人体骨架检测结果将扩充正面特征图像和/或扩充背面特征图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行二次检测,得到二次检测结果,将二次检测结果在扩充正面特征图像和/或扩充背面特征图像上进行位置映射,得到分块物品检测结果。

[0084] 在具体的实施例中,步骤S4具体包括:

[0085] S41,根据人体骨架检测结果将扩充正面特征图像和/或扩充背面特征图像划分成多个区域;

[0086] S42,基于每个区域的坐标通过RoIAlign操作对扩充正面特征图像和/或扩充背面特征图像进行采样,从而获得多个特征块,并记录每个特征块在扩充正面特征图像和/或扩充背面特征图像中的起始位置;

[0087] S43,对每个特征块采用独立的检测头进行二次检测,得到每个特征块的二次检测结果;

[0088] S44,将每个特征块的二次检测结果中的检测框坐标分别加上每个特征块在扩充正面特征图像和/或扩充背面特征图像中的起始位置,得到分块物品检测结果。

[0089] 在具体的实施例中,在步骤S43中还包括响应于第一特征块和第二特征块具有对称部位,将第二特征块进行翻转后输入第一特征块对应的检测头进行二次检测。

[0090] 具体地,本方法额外对被检测人体进行了人体骨架检测,用于区分被检测人体的四肢和躯干,然后根据各肢体部分的坐标对扩充正面特征图像和/或扩充背面特征图像进行分块,分为头颈、躯干、胯/臀部、大臂、小臂、大腿、小腿。每个特征块再经过一个独有的检测头进行检测。对于每个检测头来说,输入的背景特征相对固定,进而减少误检并提高检出率。

[0091] 以正面检测为例,根据人体骨架检测结果将正面划分为11个人体区域。参考图7,其中肢体区域8个:2-4,4-6,3-5,5-7,8-10,10-12,9-11,11-13;主干区域3个:1-2-3向上扩展区域(头),2-3-8-9(躯干),8-9(胯/臀)。同理,将关键点坐标水平翻转后,也可将背部划分为11个区域。采用RoIAlign操作,根据划分后的区域坐标对扩充正面特征图像和/或扩充背面特征图像中的特征进行采样,从而获得11个大小不等的特征块,并记录每个特征块在

扩充正面特征图像和/或扩充背面特征图像中的起始位置 (dx_i, dy_i) ,也即特征块左上角坐标。对每个特征块使用独立的检测头进行二次检测,所有对称的部位共享同一检测头,但输入特征需要水平翻转。即(2-4,3-5), (4-6,5-7), (8-10,9-11), (10-12,11-13)这几对部位共享检测头,其余部位各有一个检测头,因此共有7个不同的检测头。由此获得每个特征块独立的二次检测结果。将二次检测结果中的检测框坐标分别加上其对应的每个特征块的起始位置,从而获得二次检测结果 $[x_{\min}, y_{\min}, x_{\max}, y_{\max}]$ 在扩充正面特征图像和/或扩充背面特征图像中的位置映射: $[x_{\min}+dx_i, y_{\min}+dy_i, x_{\max}+dx_i, y_{\max}+dy_i]$ 。

[0092] S5,将分块物品检测结果与全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。

[0093] 具体地,将分块物品检测结果与全图物品检测结果合并,并进行非极大值抑制算法(NMS),获得最终的物品检测结果。

[0094] 进一步参考图8,作为对上述各图所示方法的实现,本申请提供了一种基于毫米波图像的物品检测装置的一个实施例,该装置实施例与图2所示的方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0095] 本申请实施例提供了一种基于毫米波图像的物品检测装置,包括:

[0096] 数据获取模块1,被配置为获取人体正面的毫米波图像和人体背面的毫米波图像,将人体正面的毫米波图像和人体背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征;

[0097] 特征扩充模块2,被配置为基于注意力机制将正面特征和背面特征分别进行交互融合扩充,得到扩充正面特征图像和扩充背面特征图像;

[0098] 全面检测模块3,被配置为将扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果;

[0099] 分块检测模块4,被配置为基于人体骨架检测结果将扩充正面特征图像和/或扩充背面特征图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行二次检测,得到二次检测结果,将二次检测结果在扩充正面特征图像和/或扩充背面特征图像上进行位置映射,得到分块物品检测结果;

[0100] 合并模块5,被配置为将分块物品检测结果与全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。

[0101] 下面参考图9,其示出了适于用来实现本申请实施例的电子设备的计算机装置900的结构示意图。图9示出的电子设备仅仅是一个示例,不应对本申请实施例的功能和使用范围带来任何限制。

[0102] 如图9所示,计算机装置900包括中央处理单元(CPU)901和图形处理器(GPU)902,其可以根据存储在只读存储器(ROM)903中的程序或者从存储部分909加载到随机访问存储器(RAM)904中的程序而执行各种适当的动作和处理。在RAM 904中,还存储有装置900操作所需的各种程序和数据。CPU 901、GPU902、ROM 903以及RAM904通过总线905彼此相连。输入/输出(I/O)接口906也连接至总线905。

[0103] 以下部件连接至I/O接口906:包括键盘、鼠标等的输入部分907;包括诸如、液晶显示器(LCD)等以及扬声器等的输出部分908;包括硬盘等的存储部分909;以及包括诸如LAN

卡、调制解调器等网络接口卡的通信部分910。通信部分910经由诸如因特网的网络执行通信处理。驱动器911也可以根据需要连接至I/O接口906。可拆卸介质912,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911上,以便于从其上读出的计算机程序根据需要被安装入存储部分909。

[0104] 特别地,根据本公开的实施例,上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,本公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分910从网络上被下载和安装,和/或从可拆卸介质912被安装。在该计算机程序被中央处理单元(CPU)901和图形处理器(GPU)902执行时,执行本申请的方法中限定的上述功能。

[0105] 需要说明的是,本申请所述的计算机可读介质可以是计算机可读信号介质或者计算机可读介质或者是上述两者的任意组合。计算机可读介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的装置、装置或器件,或者任意以上的组合。计算机可读介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请中,计算机可读介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行装置、装置或者器件使用或者与其结合使用。而在本申请中,计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行装置、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:无线、电线、光缆、RF等等,或者上述的任意合适的组合。

[0106] 可以以一种或多种程序设计语言或其组合来编写用于执行本申请的操作的计算机程序代码,所述程序设计语言包括面向对象的程序设计语言——诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言——诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0107] 附图中的流程图和框图,图示了按照本申请各种实施例的装置、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执

行规定的功能或操作的专用的基于硬件的装置来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0108] 描述于本申请实施例中所涉及到的模块可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的模块也可以设置在处理器中。

[0109] 作为另一方面,本申请还提供了一种计算机可读介质,该计算机可读介质可以是上述实施例中描述的设备中所包含的;也可以是单独存在,而未装配入该设备中。上述计算机可读介质承载有一个或者多个程序,当上述一个或者多个程序被该设备执行时,使得该设备:获取人体正面的毫米波图像和人体背面的毫米波图像,将人体正面的毫米波图像和人体背面的毫米波图像输入经训练的主干网络分别提取正面特征和背面特征;基于注意力机制将正面特征和背面特征分别进行交互融合扩充,得到扩充正面特征图像和扩充背面特征图像;将扩充正面特征图像和/或扩充背面特征图像输入经训练的基于非同源数据的多任务神经网络架构,输出全图物品检测结果和人体骨架检测结果;基于人体骨架检测结果将扩充正面特征图像和/或扩充背面特征图像划分成多个特征块,并对每个特征块采用独立的检测头分别进行二次检测,得到二次检测结果,将二次检测结果在扩充正面特征图像和/或扩充背面特征图像上进行位置映射,得到分块物品检测结果;将分块物品检测结果与全图物品检测结果进行非极大值抑制,得到最终的物品检测结果。

[0110] 以上描述仅为本申请的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解,本申请中所涉及的发明范围,并不限于上述技术特征的特定组合而成的技术方案,同时也应涵盖在不脱离上述发明构思的情况下,由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本申请中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。

100

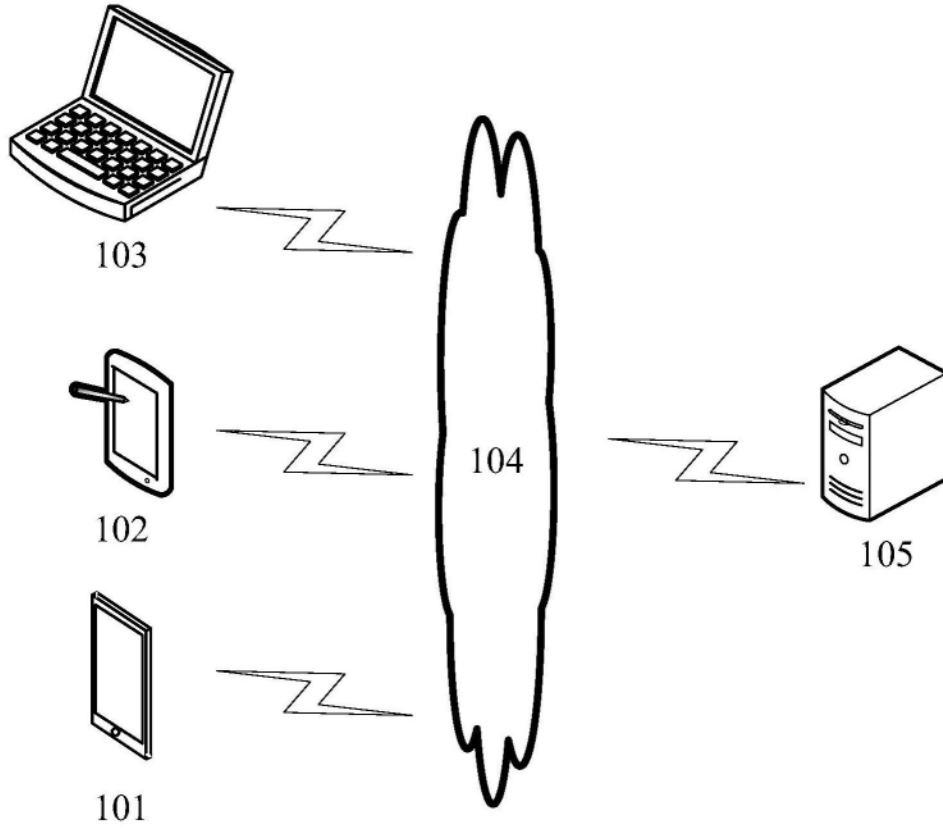


图1

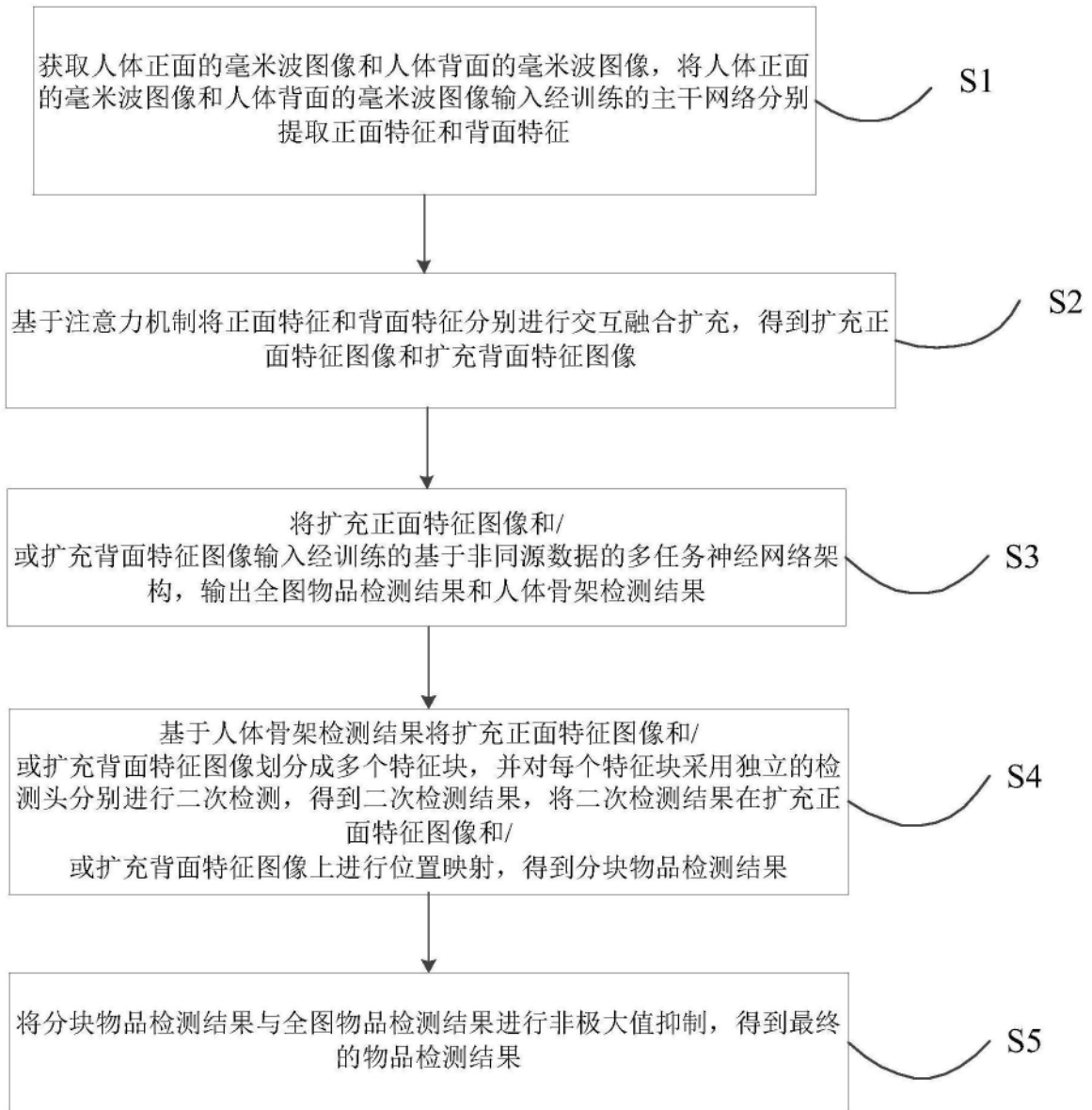


图2

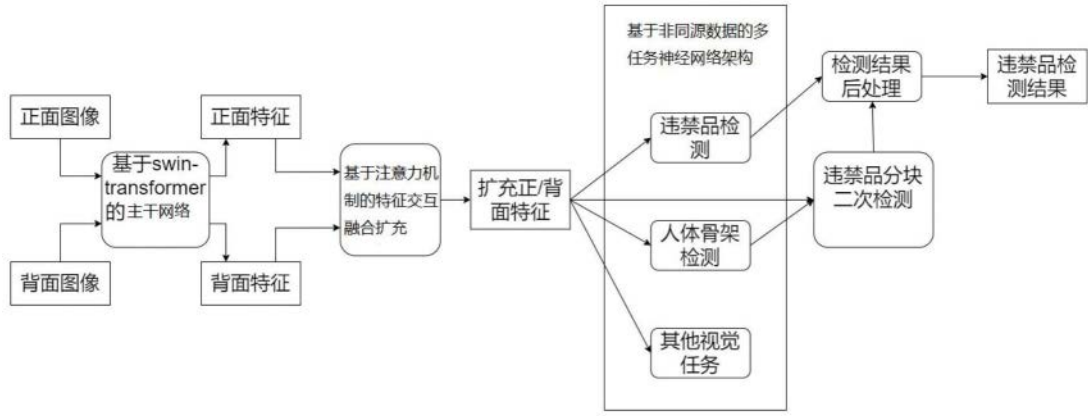


图3

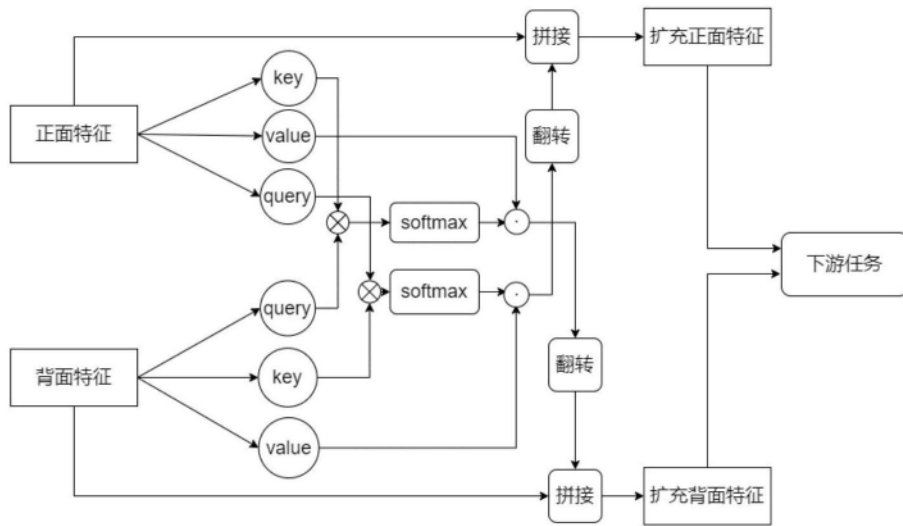


图4

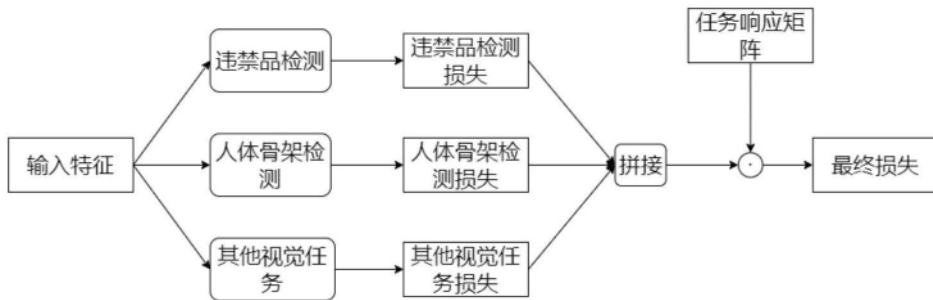


图5

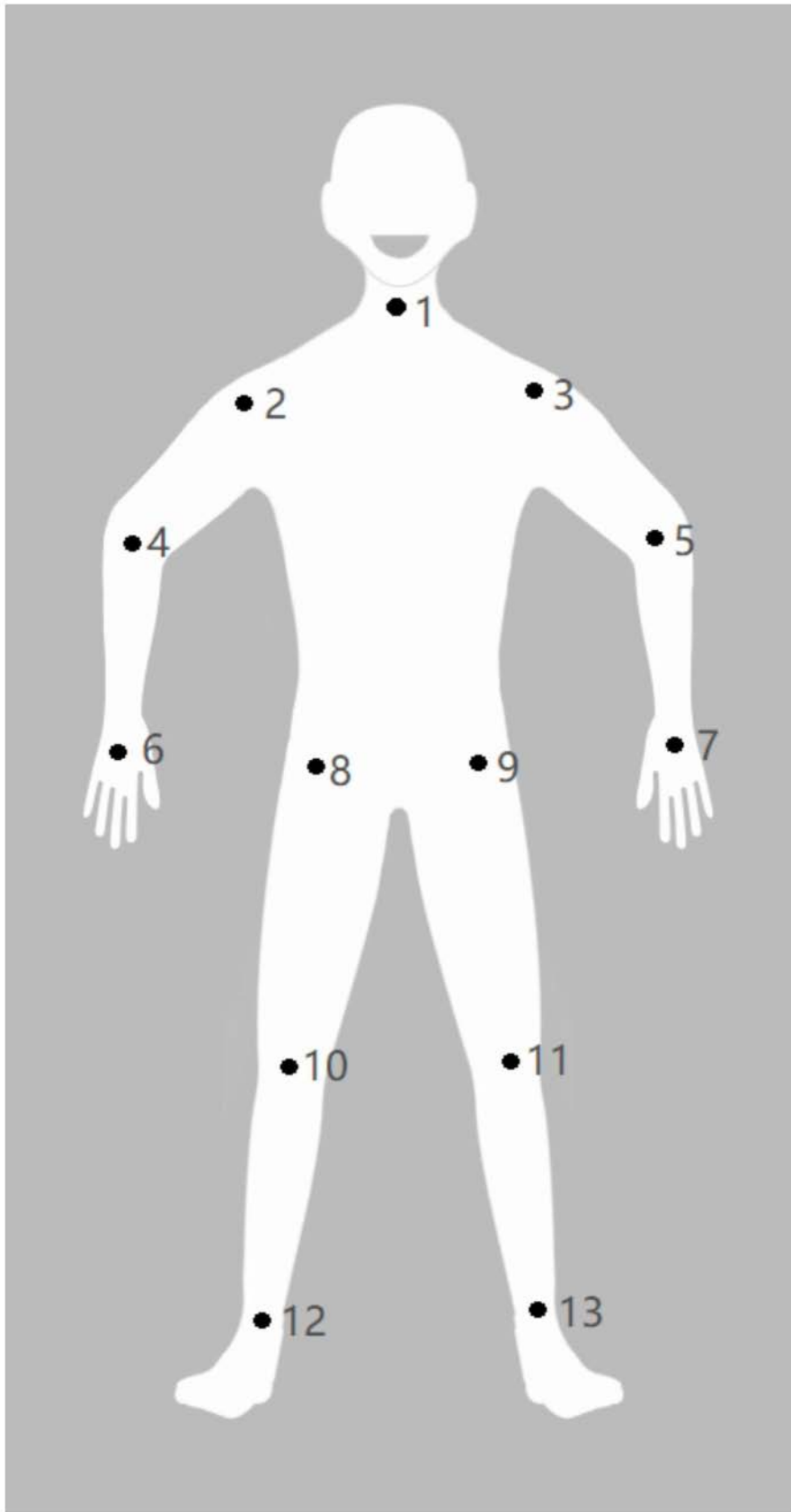


图6

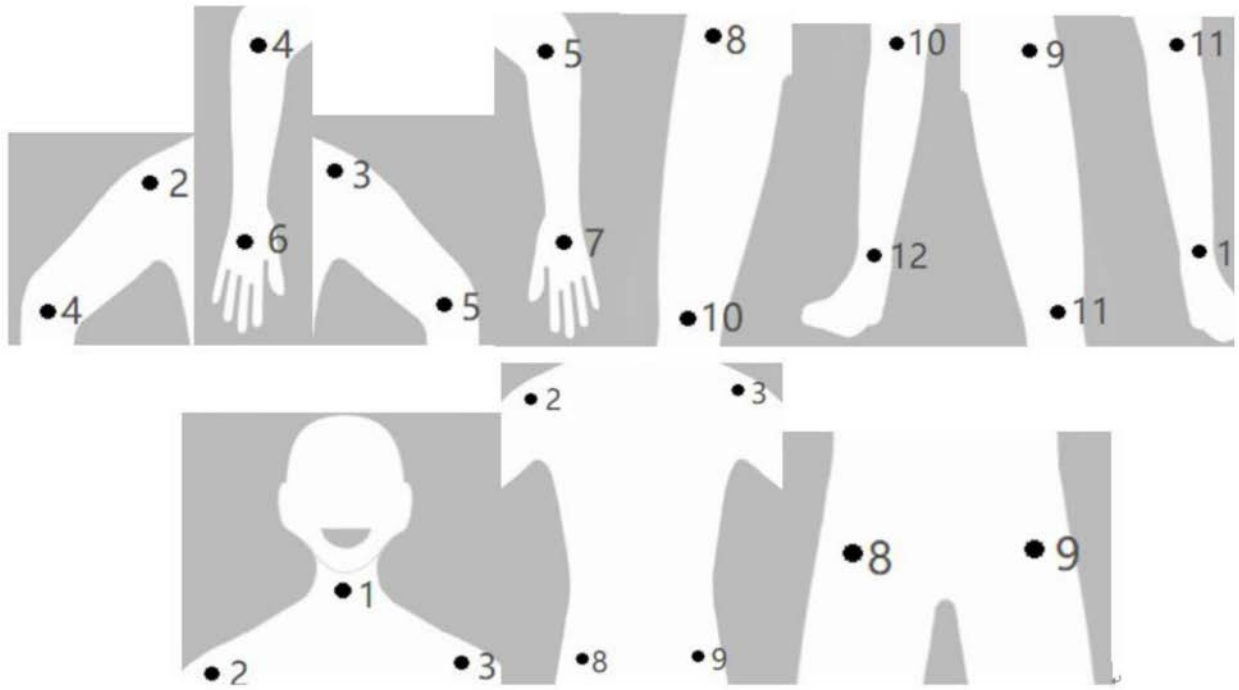


图7

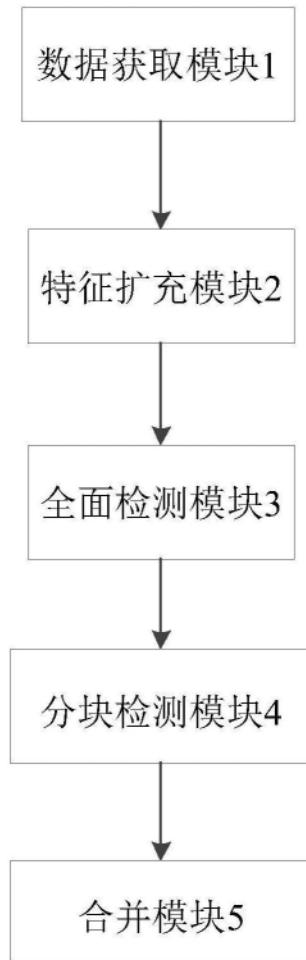


图8

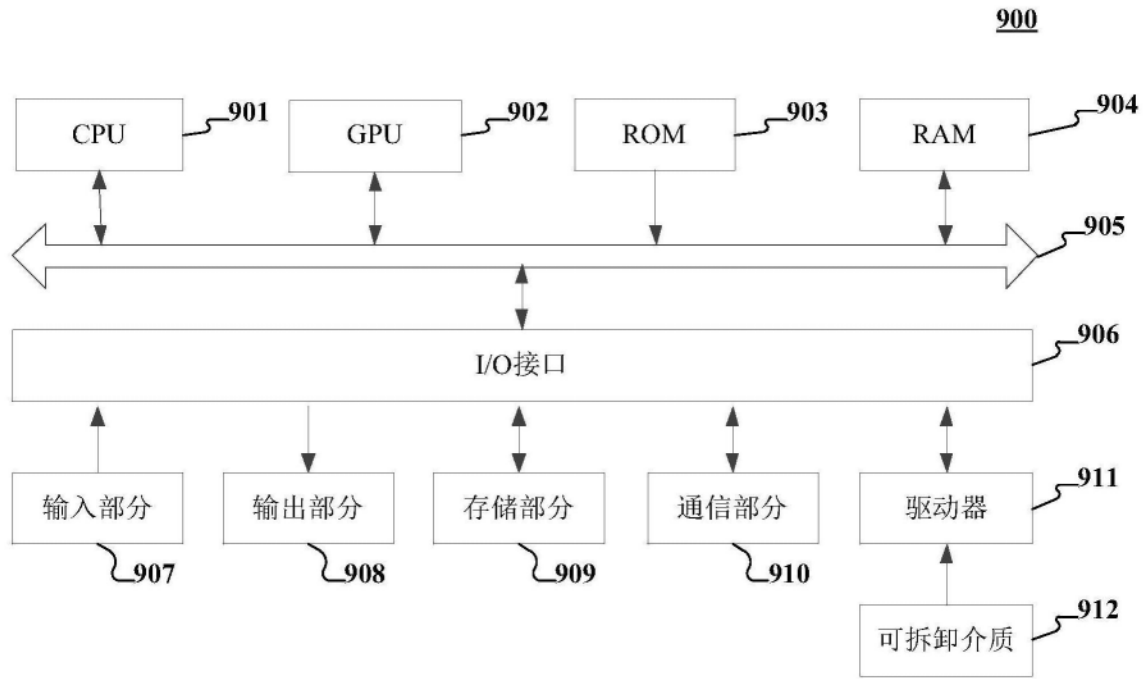


图9