



(12) 发明专利

(10) 授权公告号 CN 109801270 B

(45) 授权公告日 2021.07.16

(21) 申请号 201811639702.9

G06T 7/136 (2017.01)

(22) 申请日 2018.12.29

G06K 9/62 (2006.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 109801270 A

(56) 对比文件

CN 108764164 A, 2018.11.06

CN 108460362 A, 2018.08.28

(43) 申请公布日 2019.05.24

CN 108830196 A, 2018.11.16

(73) 专利权人 北京市商汤科技开发有限公司
地址 100084 北京市海淀区中关村东路1号
院3号楼7层710-712房间

CN 108520450 A, 2018.09.11

WO 2016086744 A1, 2016.06.09

CN 108460362 A, 2018.08.28

(72) 发明人 陈恺 王佳琦 杨硕 吕健勤
林达华

审查员 孙丹

(74) 专利代理机构 北京林达刘知识产权代理事
务所(普通合伙) 11277
代理人 刘新宇

(51) Int. Cl.

G06T 7/00 (2017.01)

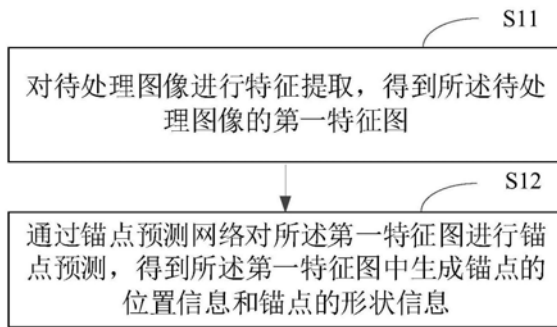
权利要求书2页 说明书12页 附图6页

(54) 发明名称

锚点确定方法及装置、电子设备和存储介质

(57) 摘要

本公开涉及一种锚点确定方法及装置、电子设备和存储介质,所述方法包括:对待处理图像进行特征提取,得到所述待处理图像的第一特征图;通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息和所述锚点的形状信息。本公开实施例能够生成非均匀且形状任意的锚点,提高物体检测的准确率。



1. 一种锚点确定方法,其特征在于,包括:
对待处理图像进行特征提取,得到所述待处理图像的第一特征图;
通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息和所述锚点的形状信息;其中,所述锚点包括非均匀的锚点,所述位置信息根据所述第一特征图中各个位置生成锚点的概率所确定。
2. 根据权利要求1所述的方法,其特征在于,所述锚点预测网络包括位置预测子网络,所述通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息,包括:
将所述第一特征图输入至所述位置预测子网络,确定所述第一特征图中各个位置生成锚点的概率;
将所述概率大于或等于预设阈值的位置信息确定为所述锚点的位置信息。
3. 根据权利要求1所述的方法,其特征在于,所述锚点预测网络包括形状预测子网络,其中,所述通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中所述锚点的形状信息,包括:
将第一特征图输入至形状预测子网络,确定所述锚点的形状信息。
4. 根据权利要求1所述的方法,其特征在于,还包括:
基于所述锚点的形状信息,通过修正网络对所述第一特征图进行特征修正,得到第二特征图。
5. 根据权利要求4所述的方法,其特征在于,所述修正网络包括第一卷积层和第二卷积层,
所述基于所述锚点的形状信息,通过修正网络对所述第一特征图进行特征修正,得到第二特征图,包括:
将所述锚点的形状信息输入至第一卷积层,得到所述第二卷积层的偏移量;
将所述第一特征图和所述偏移量输入至所述第二卷积层,得到所述第二特征图。
6. 根据权利要求4或5所述的方法,其特征在于,所述方法还包括:
对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的识别结果。
7. 根据权利要求4或5所述的方法,其特征在于,所述方法还包括:
对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的推荐区域;
对所述推荐区域和所述第二特征图进行分类和区域校正,得到对所述待处理图像的识别结果。
8. 根据权利要求1所述的方法,其特征在于,所述对待处理图像进行特征提取,得到所述待处理图像的第一特征图,包括:
利用预先训练的特征提取网络对所述待处理图像进行特征提取,得到所述待处理图像的第一特征图。
9. 一种锚点确定装置,其特征在于,包括:
特征提取模块,用于对待处理图像进行特征提取,得到所述待处理图像的第一特征图;
锚点预测模块,用于通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第

一特征图中生成锚点的位置信息和所述锚点的形状信息；其中，所述锚点包括非均匀的锚点，所述位置信息根据所述第一特征图中各个位置生成锚点的概率所确定。

10. 根据权利要求9所述的装置，其特征在于，所述锚点预测网络包括位置预测子网络，所述锚点预测模块包括：

概率生成子模块，用于将所述第一特征图输入至所述位置预测子网络，确定所述第一特征图中各个位置生成锚点的概率；

位置信息确定子模块，用于将所述概率大于或等于预设阈值的位置信息确定为所述锚点的位置信息。

11. 根据权利要求9所述的装置，其特征在于，所述锚点预测网络包括形状预测子网络，锚点预测模块包括：

形状信息确定子模块，用于将第一特征图输入至形状预测子网络，确定所述锚点的形状信息。

12. 根据权利要求9所述的装置，其特征在于，所述锚点确定装置还包括：

修正模块，用于基于所述锚点的形状信息，通过修正网络对所述第一特征图进行特征修正，得到第二特征图。

13. 根据权利要求12所述的装置，其特征在于，所述修正网络包括第一卷积层和第二卷积层，

所述修正模块包括：

第一卷积子模块，用于将所述锚点的形状信息输入至第一卷积层，得到所述第二卷积层的偏移量；

第二卷积子模块，用于将所述第一特征图和所述偏移量输入至所述第二卷积层，得到所述第二特征图。

14. 根据权利要求12或13所述的装置，其特征在于，所述装置还包括：

第一识别结果确定模块，用于对所述锚点和所述第二特征图进行分类和锚点校正，得到对所述待处理图像的分类结果。

15. 根据权利要求12或13所述的装置，其特征在于，所述装置还包括：

推荐区域确定模块，用于对所述锚点和所述第二特征图进行分类和锚点校正，得到对所述待处理图像的推荐区域；

第二识别结果确定模块，用于对所述推荐区域和所述第二特征图进行分类和区域校正，得到对所述待处理图像的分类结果。

16. 根据权利要求9所述的装置，其特征在于，所述特征提取模块包括：

特征提取子模块，用于利用预先训练的特征提取网络对所述待处理图像进行特征提取，得到所述待处理图像的第一特征图。

17. 一种电子设备，其特征在于，包括：

处理器；

用于存储处理器可执行指令的存储器；

其中，所述处理器被配置为：执行权利要求1至8中任意一项所述的方法。

18. 一种计算机可读存储介质，其上存储有计算机程序指令，其特征在于，所述计算机程序指令被处理器执行时实现权利要求1至8中任意一项所述的方法。

锚点确定方法及装置、电子设备和存储介质

技术领域

[0001] 本公开涉及图像处理技术领域,尤其涉及一种锚点确定方法及装置、电子设备和存储介质。

背景技术

[0002] 锚点是当前物体检测技术中一个基本的元素,也是区域提议算法的基石。目前的物体检测器大都依赖密集且均匀的锚点生成方案,即在全图中按照预定于的尺度,长宽比和步长来生成均匀的锚点。

发明内容

[0003] 本公开提出了一种锚点确定的技术方案。

[0004] 根据本公开的一方面,提供了一种锚点确定方法,包括:对待处理图像进行特征提取,得到所述待处理图像的第一特征图;通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息和所述锚点的形状信息。

[0005] 在一种可能的实现方式中,所述锚点预测网络包括位置预测子网络,所述通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息,包括:将所述第一特征图输入至所述位置预测子网络,确定所述第一特征图中各个位置生成锚点的概率;将所述概率大于或等于预设阈值的位置信息确定为所述锚点的位置信息。

[0006] 在一种可能的实现方式中,所述锚点预测网络包括形状预测子网络,所述通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中所述锚点的形状信息,包括:将第一特征图输入至形状预测子网络,确定所述锚点的形状信息。

[0007] 在一种可能的实现方式中,所述方法还包括基于所述锚点的形状信息,通过修正网络对所述第一特征图进行特征修正,得到第二特征图。

[0008] 在一种可能的实现方式中,所述修正网络包括第一卷积层和第二卷积层,所述基于所述锚点的形状信息,通过修正网络对所述第一特征图进行特征修正,得到第二特征图,包括:将所述锚点的形状信息输入至第一卷积层,得到所述第二卷积层的偏移量;将所述第一特征图和所述偏移量输入至所述第二卷积层,得到所述第二特征图。

[0009] 在一种可能的实现方式中,所述方法还包括:对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的分类结果。所述方法还包括:

[0010] 在一种可能的实现方式中,对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的推荐区域;对所述推荐区域和所述第二特征图进行分类和区域校正,得到对所述待处理图像的分类结果。

[0011] 在一种可能的实现方式中,所述对待处理图像进行特征提取,得到所述待处理图像的第一特征图,包括:利用预先训练的特征提取网络对所述待处理图像进行特征提取,得到所述待处理图像的第一特征图。

[0012] 根据本公开的一方面,提供了一种锚点确定装置,包括:特征提取模块,用于对待处理图像进行特征提取,得到所述待处理图像的第一特征图;锚点预测模块,用于通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息和所述锚点的形状信息。

[0013] 在一种可能的实现方式中,所述锚点预测网络包括位置预测子网络,所述锚点预测模块包括:概率生成子模块,用于将所述第一特征图输入至所述位置预测子网络,确定所述第一特征图中各个位置生成锚点的概率;位置信息确定子模块,用于将所述概率大于或等于预设阈值的位置信息确定为所述锚点的位置信息。

[0014] 在一种可能的实现方式中,所述锚点预测网络包括形状预测子网络,锚点预测模块包括:形状信息确定子模块,用于将第一特征图输入至形状预测子网络,确定所述锚点的形状信息。

[0015] 在一种可能的实现方式中,所述锚点确定装置还包括:修正模块,用于基于所述锚点的形状信息,通过修正网络对所述第一特征图进行特征修正,得到第二特征图。

[0016] 在一种可能的实现方式中,所述修正网络包括第一卷积层和第二卷积层,所述修正模块包括:第一卷积子模块,用于将所述锚点的形状信息输入至第一卷积层,得到所述第二卷积层的偏移量;第二卷积子模块,用于将所述第一特征图和所述偏移量输入至所述第二卷积层,得到所述第二特征图。

[0017] 在一种可能的实现方式中,所述装置还包括:第一识别结果确定模块,用于对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的分类结果。

[0018] 在一种可能的实现方式中,所述装置还包括:推荐区域确定模块,用于对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的分类结果;第二识别结果确定模块,用于对所述推荐区域和所述第二特征图进行分类和区域校正,得到对所述待处理图像的分类结果。

[0019] 在一种可能的实现方式中,所述特征提取模块包括:特征提取子模块,用于利用预先训练的特征提取网络对所述待处理图像进行特征提取,得到所述待处理图像的第一特征图。

[0020] 根据本公开的一方面,提供了一种电子设备,包括:处理器;用于存储处理器可执行指令的存储器;其中,所述处理器被配置为:执行上述锚点确定方法。

[0021] 根据本公开的一方面,提供了一种计算机可读存储介质,其上存储有计算机程序指令,所述计算机程序指令被处理器执行时实现上述锚点确定方法。

[0022] 根据本公开实施例的锚点确定方法,能够对待处理图像的特征图进行锚点预测,得到锚点的位置信息和锚点的形状信息,从而能够生成非均匀且大小、形状任意的锚点,提高物体检测的准确率。

[0023] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,而非限制本公开。根据下面参考附图对示例性实施例的详细说明,本公开的其它特征及方面将变得清楚。

附图说明

[0024] 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。

[0025] 图1示出根据本公开实施例的锚点确定方法的流程图。

[0026] 图2示出根据本公开实施例的锚点确定方法的流程图。

[0027] 图3示出根据本公开实施例的锚点确定方法的流程图。

[0028] 图4示出根据本公开实施例的锚点确定方法的流程图。

[0029] 图5a和图5b示出根据本公开实施例的锚点确定方法的应用场景的示意图。

[0030] 图6示出根据本公开实施例的锚点确定装置的框图。

[0031] 图7示出根据本公开实施例的一种电子设备的框图。

[0032] 图8示出根据本公开实施例的一种电子设备的框图。

具体实施方式

[0033] 以下将参考附图详细说明本公开的各种示例性实施例、特征和方面。附图中相同的附图标记表示功能相同或相似的元件。尽管在附图中示出了实施例的各种方面,但是除非特别指出,不必按比例绘制附图。

[0034] 在这里专用的词“示例性”意为“用作例子、实施例或说明性”。这里作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。

[0035] 本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,本文中术语“至少一种”表示多种中的任意一种或多种中的至少两种的任意组合,例如,包括A、B、C中的至少一种,可以表示包括从A、B和C构成的集合中选择的任意一个或多个元素。

[0036] 另外,为了更好地说明本公开,在下文的具体实施方式中给出了众多的具体细节。本领域技术人员应当理解,没有某些具体细节,本公开同样可以实施。在一些实例中,对于本领域技术人员熟知的方法、手段、元件和电路未作详细描述,以便于凸显本公开的主旨。

[0037] 图1示出根据本公开实施例的锚点确定方法的流程图。该锚点确定方法可以由终端设备或其它处理设备(例如服务器)执行,其中,终端设备可以为用户设备(User Equipment, UE)、移动设备、用户终端、终端、蜂窝电话、无绳电话、个人数字处理(Personal Digital Assistant, PDA)、手持设备、计算设备、车载设备、可穿戴设备等。在一些可能的实现方式中,该锚点确定方法可以通过处理器调用存储器中存储的计算机可读指令的方式来实现。

[0038] 如图1所示,所述方法包括:

[0039] 步骤S11,对待处理图像进行特征提取,得到所述待处理图像的第一特征图。

[0040] 在一种可能的实现方式中,所述待处理图像可以是由图像采集设备(例如摄像头)采集的某一区域(例如商场入口、道路路口等)的场景图像,也可以是直接输入的以保存的图像或者视频帧。待处理图像中可包括待检测的一个或多个对象或物体,例如人、动物、车辆等。例如,所述待处理图像可以是“金毛犬在晒太阳”场景图片,本实现方式中可以通过锚点检测“金毛犬”在场景图片中位置和形状。

[0041] 步骤S12,通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息和锚点的形状信息。

[0042] 在一种可能的实现方式中,锚点用于表示特征图中可能存在对象或物体的区域

(矩形框),其中,矩形框的中心点可用于表示锚点的位置,矩形框的长度及宽度可用于表示锚点的形状。应当理解,锚点所在的区域也可以表示为其他形状,本公开对此不作限制。

[0043] 在一种可能的实现方式中,锚点预测网络可用于预测第一特征图中可生成的锚点的位置和形状。其中,锚点预测网络可包括例如卷积神经网络CNN,本公开对锚点预测网络的具体类型不作限制。

[0044] 根据本公开实施例的锚点确定方法,能够对待处理图像的特征图进行锚点预测,得到锚点的位置信息和锚点的形状信息,从而能够生成非均匀且大小、形状任意的锚点,提高物体检测的准确率。

[0045] 在一种可能的实现方式中,可在步骤S11中对待处理图像进行特征提取,得到所述待处理图像的第一特征图。其中,步骤S11可包括:

[0046] 利用预先训练的特征提取网络对所述待处理图像进行特征提取,得到所述待处理图像的第一特征图。

[0047] 举例来说,特征提取网络可以为预先训练的卷积神经网络CNN,将待处理图像输入特征提取网络中进行处理,可以得到待处理图像的第一特征图。

[0048] 在一种可能的实现方式中,特征提取网络包括多个卷积层且各个卷积层的卷积核的尺寸不同。在该情况下,得到的第一特征图可包括尺寸不同的多个特征图。

[0049] 本公开对特征提取网络的具体网络类型以及第一特征图的数量不作限制。

[0050] 在一种可能的实现方式中,根据得到的第一特征图,可在步骤S12中通过锚点预测网络预测第一特征图中可生成锚点的位置及锚点的形状。

[0051] 其中,锚点预测网络可包括位置预测子网络,在步骤S12中通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息,可包括:

[0052] 将所述第一特征图输入至所述位置预测子网络,确定所述第一特征图中各个位置生成锚点的概率;

[0053] 将所述概率大于或等于预设阈值的位置信息确定为所述锚点的位置信息。

[0054] 举例来说,位置预测子网络可用于预测锚点的位置。位置预测子网络可至少包括卷积层和激活层(例如sigmoid激活层)。对于形状为 $M \times C \times H \times W$ 的第一特征图(其中, M 表示批尺寸(batch size), C 表示特征图的第一通道数量, H 和 W 分别表示第一特征图的高和宽),可将第一特征图输入卷积层(例如卷积核尺寸为 $A \times A$, A 为大于或等于1的整数)中处理,输出形状为 $M \times 1 \times H \times W$ 的特征图。然后将形状为 $M \times 1 \times H \times W$ 的特征图输入激活层中处理,可得到第一特征图中各个位置($H \times W$ 个位置)生成锚点的概率。

[0055] 在一种可能的实现方式中,可将概率大于或等于预设阈值的位置信息确定为锚点的位置信息。也即,可预先设定有阈值 P ($0 < P < 1$),以便过滤掉所有概率小于该阈值 P 的位置。

[0056] 其中,该阈值越大,则生成的锚点的数量越少;反之,该阈值越小,则生成的锚点的数量越多。在生成的锚点的数量较少时,推荐区域较小,目标检测的运算量也较少。本领域技术人员可根据实际情况设定阈值 P ,本公开对阈值 P 的具体取值不作限制。

[0057] 通过这种方式,可以确定第一特征图中可生成的锚点的位置信息,从而生成非均匀的锚点。

[0058] 在一种可能的实现方式中,锚点预测网络可包括形状预测子网络,在步骤S12中通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中所述锚点的形状

信息,可包括:

[0059] 将第一特征图输入至形状预测子网络,确定所述锚点的形状信息。

[0060] 举例来说,形状预测子网络可用于预测锚点的形状。形状预测子网络可至少包括卷积层。对于形状为 $M \times C \times H \times W$ 的第一特征图,可将第一特征图输入该卷积层(例如卷积核尺寸为 $B \times B$, B 为大于或等于1的整数)中处理,输出形状为 $M \times 2 \times H \times W$ 的特征图。其中,2表示两个通道,分别表示每个位置预测的锚点高度和宽度(相对值),记为 dh 和 dw 。

[0061] 在一种可能的实现方式中,锚点高度的实际值 h 可以通过公式 $h = 8 * s * \exp(dh)$ 计算得到,锚点宽度的实际值 w 可以通过公式 $w = 8 * s * \exp(dw)$ 计算得到,其中 s 用于表示形状预测子网络的卷积层的步长。

[0062] 在一种可能的实现方式中,根据预测的每个位置的锚点高度和宽度,可确定第一特征图中所有锚点的形状信息。这样,根据位置预测子网络得到的锚点位置信息,就可以确定出可生成的锚点的形状信息。

[0063] 通过这种方式,可以确定第一特征图中可生成的锚点的形状信息,从而生成大小、形状任意的锚点。

[0064] 在一种可能的实现方式中,可通过四维特征向量 (x, y, a, b) ,来表示锚点的位置信息和形状信息。其中, x 和 y 分别表示锚点的位置(矩形框的中心点)在第一特征图中的横纵坐标, a 和 b 分别表示锚点的区域(矩形框)的高度和宽度。

[0065] 在一种可能的实现方式中,还可通过四维特征向量 (x_1, y_1, x_2, y_2) ,来表示锚点的位置信息和形状信息。其中, x_1 和 y_1 分别表示锚点的区域(矩形框)的左上顶点在第一特征图中的横纵坐标, x_2 和 y_2 分别表示锚点的区域的右下顶点在第一特征图中的横纵坐标。本公开对锚点的位置信息和形状信息的具体表示方式不作限制。

[0066] 图2示出根据本公开实施例的锚点确定方法的流程图。在一种可能的实现方式中,如图2所示,所述锚点确定方法还可包括:

[0067] 步骤S13、基于所述锚点的形状信息,通过修正网络对所述第一特征图进行特征修正,得到第二特征图。

[0068] 在一种可能的实现方式中,修正网络可用于修正第一特征图的偏差。将锚点的形状信息输入修正网络中处理,可输出修正后的特征图(第二特征图)。其中,修正网络可以例如为卷积神经网络CNN,本公开对修正网络的具体网络类型不作限制。

[0069] 在一种可能的实现方式中,所述修正网络可包括第一卷积层和第二卷积层,步骤S13可包括:

[0070] 将所述锚点的形状信息输入至第一卷积层,得到所述第二卷积层的偏移量;

[0071] 将所述第一特征图和所述偏移量输入至所述第二卷积层,得到所述第二特征图。

[0072] 举例来说,修正网络的第一卷积层可用于确定第二卷积层的偏移量,其中,第一卷积层的卷积核尺寸可例如为 1×1 。将锚点的形状信息输入至第一卷积层,可输出第二卷积层的偏移量。

[0073] 在一种可能的实现方式中,修正网络的第二卷积层可用于对特征图进行修正。其中,第二卷积层可以为可变形卷积层,其卷积核尺寸可为 $N \times N$ (N 为大于1的整数,例如 $N = 3$)。将第一特征图和偏移量输入至第二卷积层中处理,可输出修正后的特征图(第二特征图)。

[0074] 通过这种方式,可以实现对第一特征图的修正,从而提高物体检测的准确率。

[0075] 根据本公开实施例的锚点确定方法可以用于物体检测,例如可以应用于单阶段(single-stage)物体检测或二阶段(two-stage)物体检测。

[0076] 图3示出根据本公开实施例的锚点确定方法的流程图。在一种可能的实现方式中,如图3所示,所述锚点确定方法还可包括:

[0077] 步骤S14,对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的识别结果。

[0078] 举例来说,对于单阶段的物体检测,可直接对锚点和第二特征图进行分类和锚点校正,得到识别结果。其中,分类和锚点校正可以同时执行也可以先后执行,本公开对分类和锚点校正的执行顺序不作限制。

[0079] 在一种可能的实现方式中,对于锚点的区域(矩形框)的分类,可将锚点和第二特征图输入到分类网络中处理,该分类网络可包括一个或多个卷积层以及输出通道数为类别数的卷积层。在经过该分类网络的一个或多个卷积层之后,再经过输出通道数为类别数的卷积层,可输出锚点属于每个类别的分类概率。这样,可将锚点属于各个类别的分类概率确定为锚点分类结果。

[0080] 在一种可能的实现方式中,对于锚点校正,可通过锚点边框回归的方式实现校正过程。可将锚点和第二特征图输入到锚点校正网络中处理,该锚点校正网络可包括一个或多个卷积层以及输出通道数为4的卷积层。在经过该锚点校正网络的一个或多个卷积层之后,再经过输出通道数为4的卷积层,可输出锚点的区域(矩形框)的位置(四维特征向量),例如矩形框的中心点坐标和宽高,或矩形框的左上顶点坐标和右下顶点坐标等。这样,可将矩形框的位置确定为锚点校正结果。

[0081] 在一种可能的实现方式中,根据得到的锚点分类结果和锚点校正结果,可确定待处理图像的识别结果。也即识别结果包括检测到的待处理图像中的物体的类别和区域坐标。

[0082] 通过这种方式,可以得到单阶段物体检测的识别结果,提高物体检测的准确率。

[0083] 图4示出根据本公开实施例的锚点确定方法的流程图。在一种可能的实现方式中,如图4所示,所述锚点确定方法还可包括:

[0084] 步骤S15,对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的推荐区域;

[0085] 步骤S16,对所述推荐区域和所述第二特征图进行分类和区域校正,得到对所述待处理图像的识别结果。

[0086] 举例来说,对于二阶段的物体检测,可先在步骤S15中对锚点和第二特征图进行分类和锚点校正,得到待处理图像的推荐区域。其中,分类和锚点校正可以同时执行也可以先后执行,本公开对分类和锚点校正的执行顺序不作限制。

[0087] 在一种可能的实现方式中,对于步骤S15中的分类,可将锚点和第二特征图输入到分类卷积网络(例如二分类器)中处理,得到各个锚点的类别,并将预设类别(例如前景类别)的锚点的区域确定为待处理图像的推荐区域。本公开对分类卷积网络的具体类型不作限制。

[0088] 在一种可能的实现方式中,对于步骤S15中的锚点校正,可通过锚点边框回归的方

式实现校正过程。可将锚点和第二特征图输入到锚点校正网络中处理,可输出所有锚点的区域(矩形框)的位置。这样,根据分类得到的推荐区域,可确定出锚点校正后的推荐区域的位置。本公开对锚点校正网络的具体类型不作限制。

[0089] 在一种可能的实现方式中,在确定推荐区域之后,可在步骤S16中对所述推荐区域和所述第二特征图进行分类和区域校正,得到对所述待处理图像的认识结果。其中,分类和区域校正可以同时进行也可以先后执行,本公开对分类和区域校正的执行顺序不作限制。

[0090] 在一种可能的实现方式中,对于步骤S16中的分类,可将推荐区域和第二特征图输入到分类网络(例如多分类器)中处理,输出推荐区域属于每个类别的分类概率。这样,可将推荐区域属于各个类别的分类概率确定为推荐区域的分类结果。本公开对分类网络的具体类型不作限制。

[0091] 在一种可能的实现方式中,对于步骤S16中的区域校正,可通过锚点边框回归的方式实现校正过程。可将推荐区域和第二特征图输入到区域校正网络中处理,输出推荐区域的位置。这样,可确定推荐区域的校正结果。

[0092] 在一种可能的实现方式中,根据推荐区域的分类结果和校正结果,可确定待处理图像的认识结果。也即认识结果包括检测到的待处理图像中的物体的类别和区域坐标。

[0093] 通过这种方式,可以得到二阶段物体检测的认识结果,提高物体检测的准确率。

[0094] 在一种可能的实现方式中,所述锚点确定方法还可包括对锚点预测网络的训练过程。例如通过已标注的训练集来训练所述锚点预测网络。训练集的标注信息用于标注训练集的训练图像中对象/物体的位置、形状和/或类别。

[0095] 在一种可能的实现方式中,训练所述锚点预测网络的过程中,可以将训练集中的训练图像输入所述锚点预测网络中处理,获得预测结果;根据预测结果及训练图像的标注信息,确定锚点预测网络的网络损失;根据所述网络损失,训练锚点预测网络。

[0096] 其中,所述锚点预测网络的网络损失可以包括位置预测子网络和形状预测子网络的网络损失。锚点预测网络的网络损失可以表示为:

$$[0097] \quad L = \lambda_1 L_{loc} + \lambda_2 L_{shape} + L_{cls} + L_{reg}$$

[0098] 其中, L 用于表示锚点预测网络的网络损失, L_{loc} 用于表示位置预测子网络的网络损失, L_{shape} 用于表示形状预测子网络的网络损失, L_{cls} 用于表示锚点预测网络的回归损失, L_{reg} 用于表示锚点预测网络的分类损失。 λ_1 用于表示位置预测子网络的网络损失的权重, λ_2 用于表示形状预测子网络的网络损失权重。

[0099] 本公开对锚点预测网络的具体训练方式不作限制。

[0100] 图5a和图5b示出根据本公开实施例的锚点确定方法的应用场景的示意图。如图5a所示,在该实现方式中,可从包括多个第一特征图的特征集510中获取多个第一特征图,并行输入至各个网络中,每个锚点预测网络会分别得到锚点(anchors)的位置和形状信息520和修正后的特征图(第二特征图530)。基于锚点的位置和形状信息520和第二特征图530,可进行单阶段物体检测或二阶段物体检测,得到认识结果(prediction)。

[0101] 如图5b所示,在该实现方式中,该网络包括锚点预测网络51和修正网络52,其中锚点预测网络51包括位置预测子网络和形状预测子网络。

[0102] 在该实现方式中,可将待处理图像进行特征提取得到第一特征图510;将第一特征图510输入锚点预测网络51的位置预测子网络和形状预测子网络中处理,可输出锚点的位

置和形状信息520;将第一特征图510以及锚点的形状信息输入修正网络52中处理,可输出修正后的特征图(第二特征图530),从而实现锚点确定及特征图修正的整个过程。

[0103] 根据本公开实施例的锚点确定方法,能够取代相关技术中采用滑窗方式生成的均匀的、形状预先定义的锚点,生成稀疏的、任意形状的锚点,并根据锚点的形状来修正特征图。根据本公开的实施例,能够显著提高区域提议网络(区域推荐)的召回率;能够应用在人脸检测、自动驾驶等场景下,提高人脸、车辆或一般物体等的检测准确率以及检测速度。

[0104] 可以理解,本公开提及的上述各个方法实施例,在不违背原理逻辑的情况下,均可以彼此相互结合形成结合后的实施例,限于篇幅,本公开不再赘述。

[0105] 图6示出根据本公开实施例的锚点确定装置的框图,如图6所示,所述锚点确定装置60包括:

[0106] 特征提取模块61,用于对待处理图像进行特征提取,得到所述待处理图像的第一特征图;

[0107] 锚点预测模块62,用于通过锚点预测网络对所述第一特征图进行锚点预测,得到所述第一特征图中生成锚点的位置信息和所述锚点的形状信息。

[0108] 在一种可能的实现方式中,所述锚点预测网络包括位置预测子网络,所述锚点预测模块包括:概率生成子模块,用于将所述第一特征图输入至所述位置预测子网络,确定所述第一特征图中各个位置生成锚点的概率;位置信息确定子模块,用于将所述概率大于或等于预设阈值的位置信息确定为所述锚点的位置信息。

[0109] 在一种可能的实现方式中,所述锚点预测网络包括形状预测子网络,锚点预测模块包括:形状信息确定子模块,用于将第一特征图输入至形状预测子网络,确定所述锚点的形状信息。

[0110] 在一种可能的实现方式中,所述锚点确定装置还包括:修正模块,用于基于所述锚点的形状信息,通过修正网络对所述第一特征图进行特征修正,得到第二特征图。

[0111] 在一种可能的实现方式中,所述修正网络包括第一卷积层和第二卷积层,所述修正模块包括:第一卷积子模块,用于将所述锚点的形状信息输入至第一卷积层,得到所述第二卷积层的偏移量;第二卷积子模块,用于将所述第一特征图和所述偏移量输入至所述第二卷积层,得到所述第二特征图。

[0112] 在一种可能的实现方式中,所述装置还包括:第一识别结果确定模块,用于对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的识别结果。

[0113] 在一种可能的实现方式中,所述装置还包括:推荐区域确定模块,用于对所述锚点和所述第二特征图进行分类和锚点校正,得到对所述待处理图像的推荐区域;第二识别结果确定模块,用于对所述推荐区域和所述第二特征图进行分类和区域校正,得到对所述待处理图像的识别结果。

[0114] 在一种可能的实现方式中,所述特征提取模块包括:特征提取子模块,用于利用预先训练的特征提取网络对所述待处理图像进行特征提取,得到所述待处理图像的第一特征图。

[0115] 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。

[0116] 本公开实施例还提出一种计算机可读存储介质,其上存储有计算机程序指令,所述计算机程序指令被处理器执行时实现上述方法。计算机可读存储介质可以是非易失性计算机可读存储介质。

[0117] 本公开实施例还提出一种电子设备,包括:处理器;用于存储处理器可执行指令的存储器;其中,所述处理器被配置为上述锚点确定方法。

[0118] 图7是根据一示例性实施例示出的一种电子设备800的框图。例如,电子设备800可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等终端。

[0119] 参照图7,电子设备800可以包括以下一个或多个组件:处理组件802,存储器804,电源组件806,多媒体组件808,音频组件810,输入/输出(I/O)的接口812,传感器组件814,以及通信组件816。

[0120] 处理组件802通常控制电子设备800的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理组件802可以包括一个或多个处理器820来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件802可以包括一个或多个模块,便于处理组件802和其他组件之间的交互。例如,处理组件802可以包括多媒体模块,以方便多媒体组件808和处理组件802之间的交互。

[0121] 存储器804被配置为存储各种类型的数据以支持在电子设备800的操作。这些数据的示例包括用于在电子设备800上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器804可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0122] 电源组件806为电子设备800的各种组件提供电力。电源组件806可以包括电源管理系统,一个或多个电源,及其他与为电子设备800生成、管理和分配电力相关联的组件。

[0123] 多媒体组件808包括在所述电子设备800和用户之间提供一个输出接口的屏幕。在一些实施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件808包括一个前置摄像头和/或后置摄像头。当电子设备800处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0124] 音频组件810被配置为输出和/或输入音频信号。例如,音频组件810包括一个麦克风(MIC),当电子设备800处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器804或经由通信组件816发送。在一些实施例中,音频组件810还包括一个扬声器,用于输出音频信号。

[0125] I/O接口812为处理组件802和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0126] 传感器组件814包括一个或多个传感器,用于为电子设备800提供各个方面的状态评估。例如,传感器组件814可以检测到电子设备800的打开/关闭状态,组件的相对定位,例如所述组件为电子设备800的显示器和小键盘,传感器组件814还可以检测电子设备800或电子设备800一个组件的位置改变,用户与电子设备800接触的存在或不存在,电子设备800方位或加速/减速和电子设备800的温度变化。传感器组件814可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件814还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件814还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0127] 通信组件816被配置为便于电子设备800和其他设备之间有线或无线方式的通信。电子设备800可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信组件816经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信组件816还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0128] 在示例性实施例中,电子设备800可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0129] 在示例性实施例中,还提供了一种非易失性计算机可读存储介质,例如包括计算机程序指令的存储器804,上述计算机程序指令可由电子设备800的处理器820执行以完成上述方法。

[0130] 图8是根据一示例性实施例示出的一种电子设备1900的框图。例如,电子设备1900可以被提供为一服务器。参照图8,电子设备1900包括处理组件1922,其进一步包括一个或多个处理器,以及由存储器1932所代表的存储器资源,用于存储可由处理组件1922的执行的指令,例如应用程序。存储器1932中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外,处理组件1922被配置为执行指令,以执行上述方法。

[0131] 电子设备1900还可以包括一个电源组件1926被配置为执行电子设备1900的电源管理,一个有线或无线网络接口1950被配置为将电子设备1900连接到网络,和一个输入输出(I/O)接口1958。电子设备1900可以操作基于存储在存储器1932的操作系统,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™或类似。

[0132] 在示例性实施例中,还提供了一种非易失性计算机可读存储介质,例如包括计算机程序指令的存储器1932,上述计算机程序指令可由电子设备1900的处理组件1922执行以完成上述方法。

[0133] 本公开可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质,其上载有用于使处理器实现本公开的各个方面的计算机可读程序指令。

[0134] 计算机可读存储介质可以是保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是一—但不限于—电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、静态随机存取存储器(SRAM)、便携式

压缩盘只读存储器 (CD-ROM)、数字多功能盘 (DVD)、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身,诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波(例如,通过光纤电缆的光脉冲)、或者通过电线传输的电信号。

[0135] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备,或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令,并转发该计算机可读程序指令,以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0136] 用于执行本公开操作的计算机程序指令可以是汇编指令、指令集架构 (ISA) 指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码,所述编程语言包括面向对象的编程语言—诸如 Smalltalk、C++ 等,以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令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)。在一些实施例中,通过利用计算机可读程序指令的状态信息来个性化定制电子电路,例如可编程逻辑电路、现场可编程门阵列 (FPGA) 或可编程逻辑阵列 (PLA),该电子电路可以执行计算机可读程序指令,从而实现本公开的各个方面。

[0137] 这里参照根据本公开实施例的方法、装置(系统)和计算机程序产品的流程图和/或框图描述了本公开的各个方面。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机可读程序指令实现。

[0138] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器,从而生产出一种机器,使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时,产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中,这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作,从而,存储有指令的计算机可读介质则包括一个制品,其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

[0139] 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上,使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤,以产生计算机实现的过程,从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

[0140] 附图中的流程图和框图显示了根据本公开的多个实施例的系统、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、程序段或指令的一部分,所述模块、程序段或指令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。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,两个连续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。也要注意的,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动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0141] 以上已经描述了本公开的各实施例,上述说明是示例性的,并非穷尽性的,并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下,对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择,旨在最好地解释各实施例的原理、实际应用或对市场中的技术改进,或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。

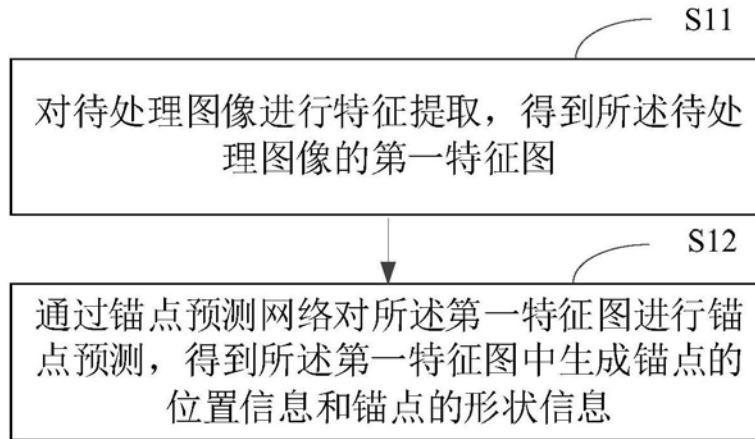


图1

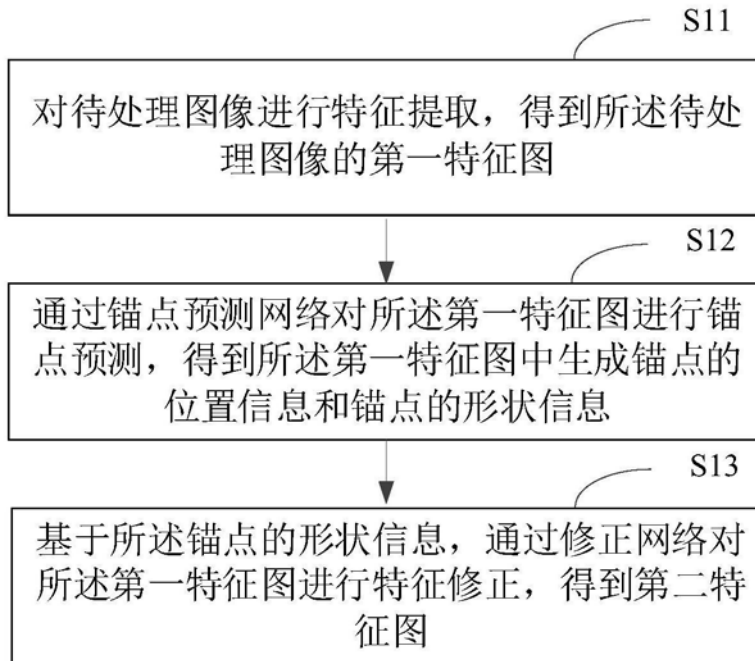


图2

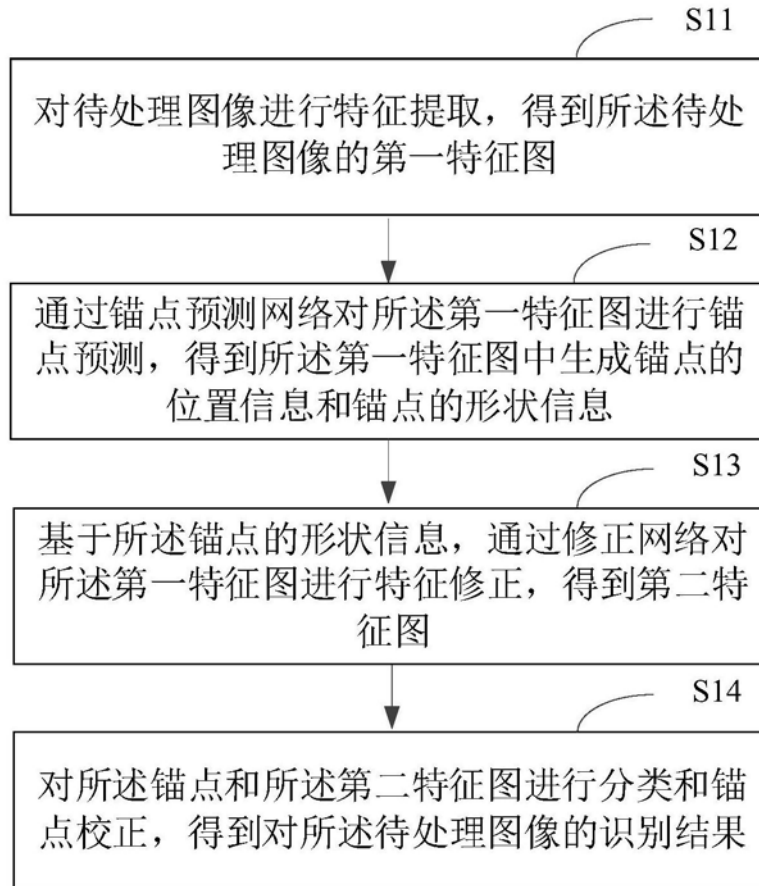


图3

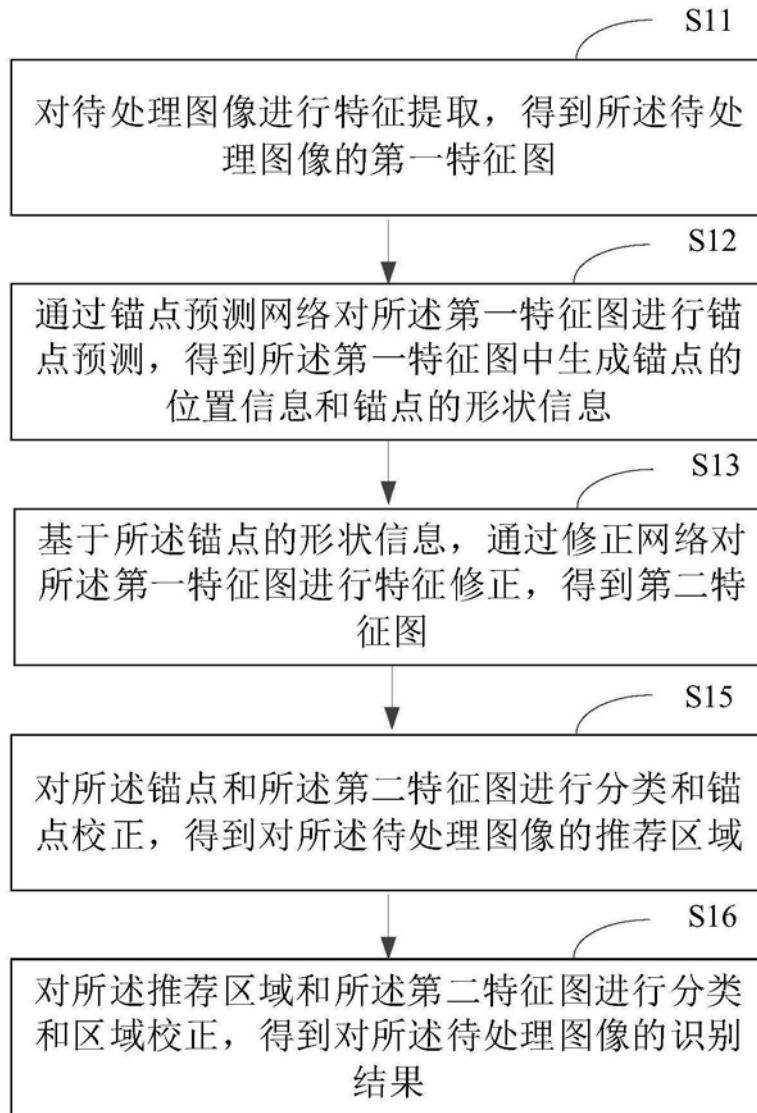


图4

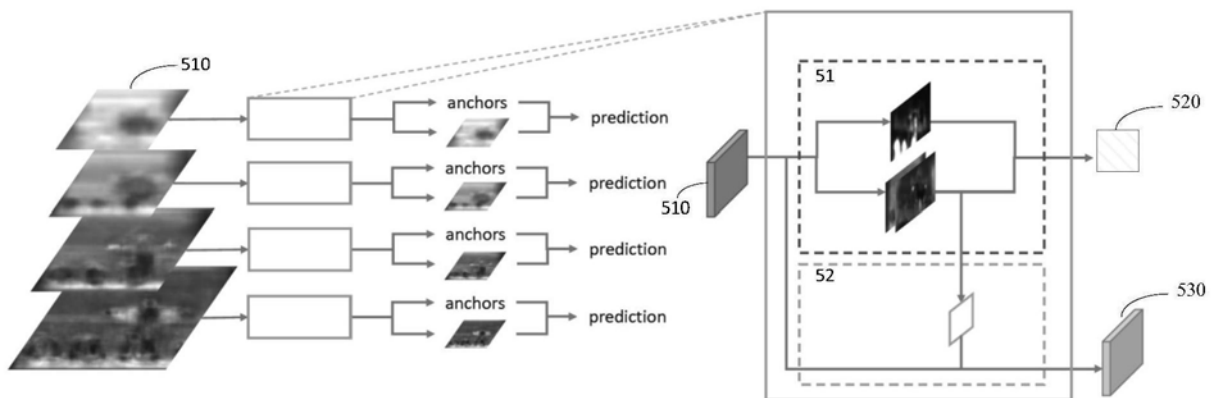


图5a

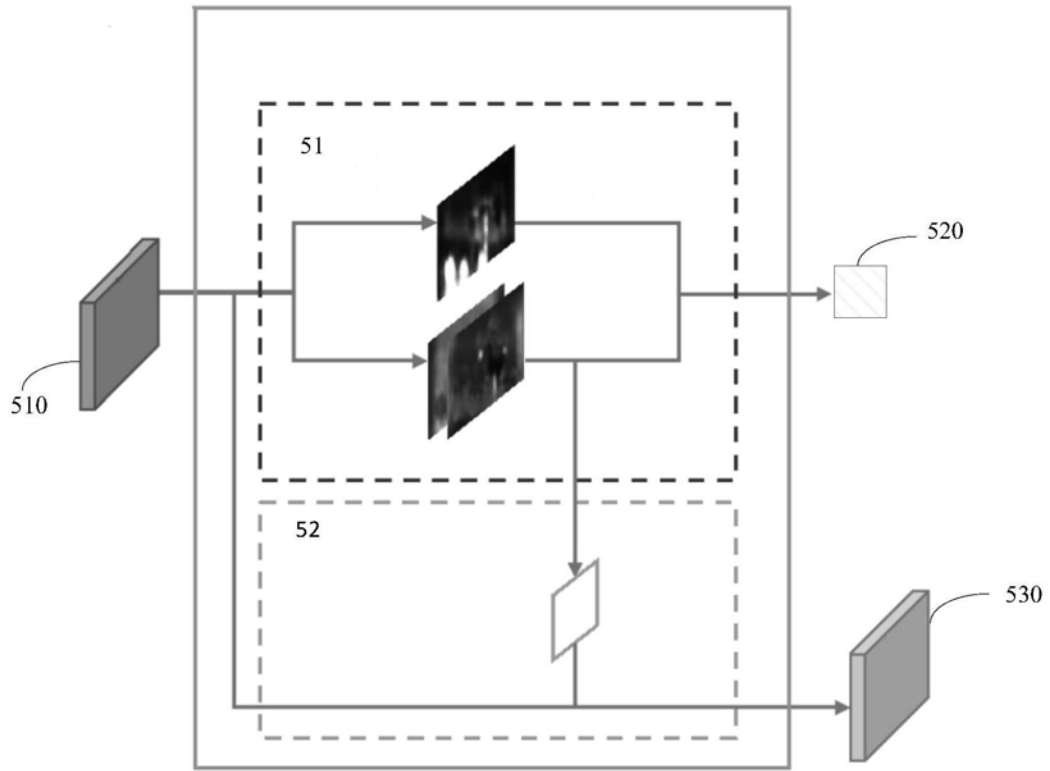


图5b

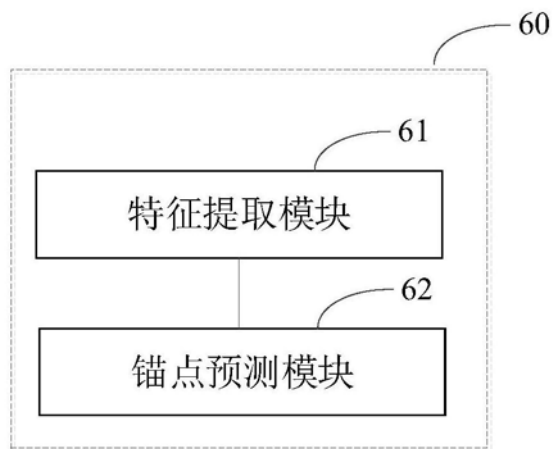


图6

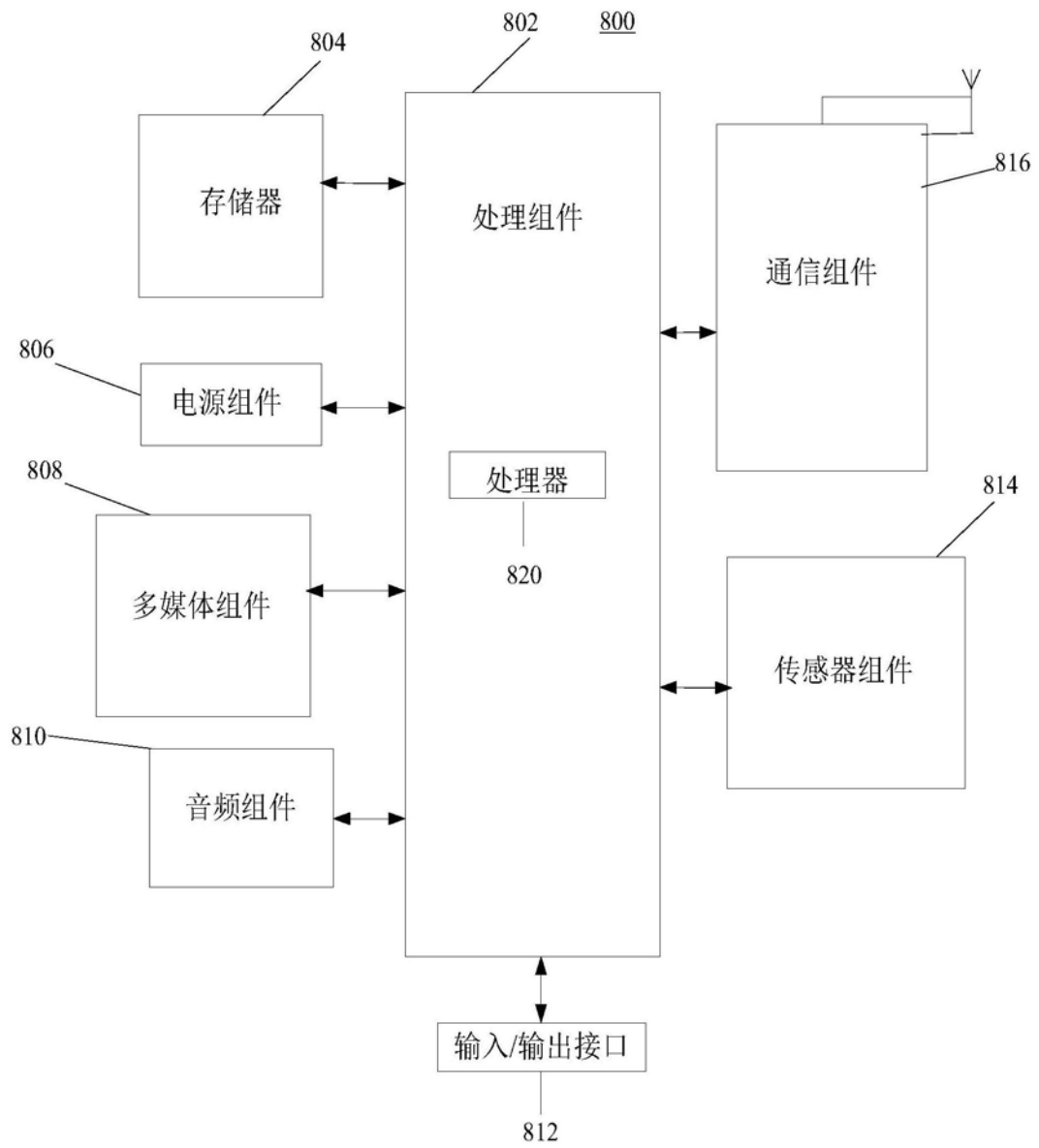


图7

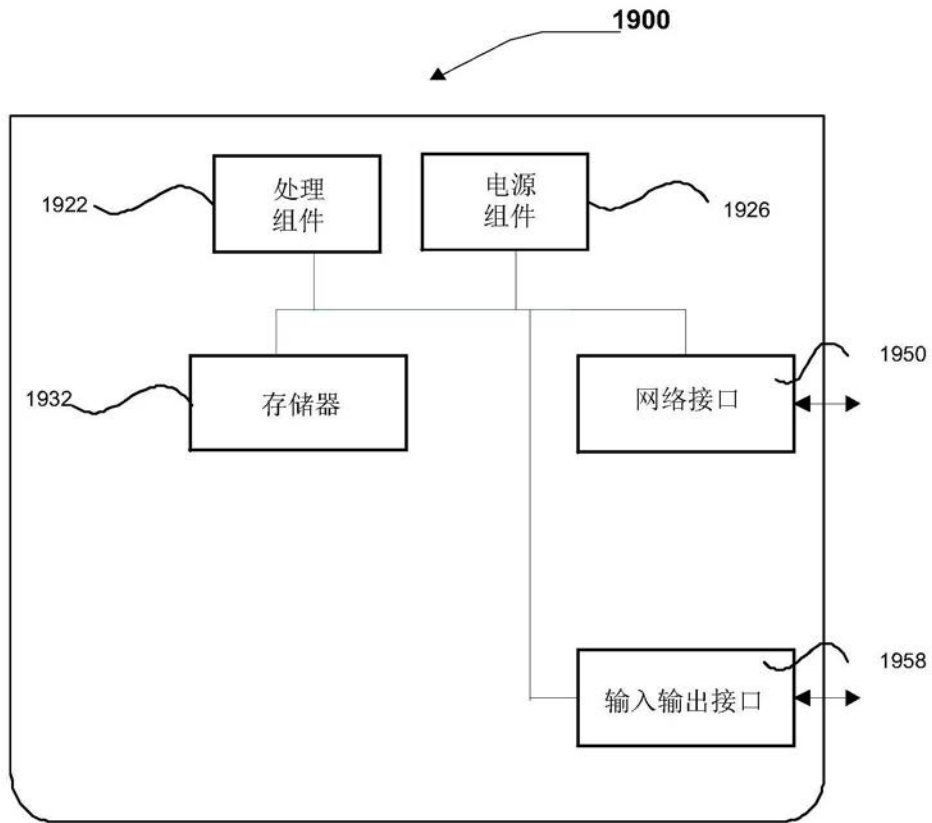


图8