

(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 8 月 11 日 (11.08.2022)



(10) 国际公布号
WO 2022/166293 A1

(51) 国际专利分类号:
G06K 9/62 (2022.01)

山东省潍坊市高新技术产业开发区东方路268号, Shandong 261031 (CN)。

(21) 国际申请号: PCT/CN2021/130102

(22) 国际申请日: 2021 年 11 月 11 日 (11.11.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110148538.7 2021年2月3日 (03.02.2021) CN

(71) 申请人: 歌尔股份有限公司(GOERTEK INC.) [CN/CN]; 中国山东省潍坊市高新技术产业开发区东方路268号, Shandong 261031 (CN)。

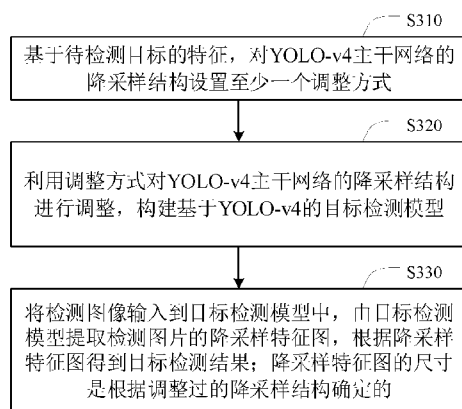
(72) 发明人: 张一凡 (ZHANG, Yifan); 中国山东省潍坊市高新技术产业开发区东方路268号, Shandong 261031 (CN)。 刘杰 (LIU, Jie); 中国

(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, IT, 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,

(54) **Title:** TARGET DETECTION METHOD AND APPARATUS

(54) 发明名称: 一种目标检测方法和装置



S310 On the basis of features of a target to be detected, set at least one adjustment mode for a down-sampling structure of a YOLO-v4 backbone network
S320 Adjust the down-sampling structure of the YOLO-v4 backbone network by using the adjustment mode, so as to construct a YOLO-v4-based target detection model
S330 Input a detection image into the target detection model, and extract a down-sampling feature map of the detection image by means of the target detection model, so as to obtain a target detection result according to the down-sampling feature map, wherein the size of the down-sampling feature map is determined according to the adjusted down-sampling structure

图 3

(57) **Abstract:** Disclosed are a target detection method and apparatus. The method comprises: on the basis of features of a target to be detected, setting at least one adjustment mode for a down-sampling structure of a YOLO-v4 backbone network; adjusting the down-sampling structure of the YOLO-v4 backbone network by using the adjustment mode, so as to construct a YOLO-v4-based target detection model; and inputting a detection image into the target detection model, and extracting a down-sampling feature map of the detection image by means of the target detection model, so as to obtain a target detection result according to the down-sampling feature map, wherein the size of the down-sampling feature map is determined according to the adjusted down-sampling structure.

(57) **摘要:** 本申请公开了一种目标检测方法和装置。该方法包括: 基于待检测目标的特征, 对YOLO-v4主干网络的降采样结构设置至少一个调整方式; 利用调整方式对YOLO-v4主干网络的降采样结构进行调整, 构建基于YOLO-v4的目标检测模型; 将检测图像输入到目标检测模型中, 由目标检测模型提取检测图像的降采样特征图, 根据降采样特征图得到目标检测结果; 降采样特征图的尺寸是根据调整过的降采样结构确定的。

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))。

一种目标检测方法和装置

技术领域

本申请涉及计算机视觉技术领域，尤其涉及一种目标检测方法和装置。

发明背景

YOLO（英文全称为 You Only Look Once）是典型的单阶段目标检测技术，即直接根据原始图像回归出目标的位置和类别等信息，目前已经发展到第四个版本，即 YOLO-v4。图 1 示出了 YOLO-v4 的网络结构示意图，可以看出其包含由多个降采样层组成的降采样结构，但这种设置存在着一些缺点，例如在工业上的缺陷检测场景，有一部分缺陷还是难以准确识别，该技术仍存在改进空间。

需要说明的是，这里的陈述仅提供与本申请有关的背景信息，而不必然地构成现有技术。

发明内容

本申请实施例提供了一种目标检测方法和装置，以进一步提高目标检测的精度。

本申请实施例采用下述技术方案：

第一方面，本申请实施例提供一种目标检测方法，包括：基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式；利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整，构建基于 YOLO-v4 的目标检测模型；将检测图像输入到目标检测模型中，由目标检测模型提取检测图像的降采样特征图，根据降采样特征图得到目标检测结果；降采样特征图的尺寸是根据调整过的降采样结构确定的。

第二方面，本申请实施例还提供一种目标检测装置，包括：调整单元，用于基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式；构建单元，用于利用调整方式对 YOLO-v4 主干网络的降采样结构进行

调整，构建基于 YOLO-v4 的目标检测模型；检测单元，用于将检测图像输入到目标检测模型中，由目标检测模型提取检测图像的降采样特征图，根据降采样特征图得到目标检测结果；降采样特征图的尺寸是根据调整过的降采样结构确定的。

第三方面，本申请实施例还提供一种电子设备，包括：处理器；以及被安排成存储计算机可执行指令的存储器，可执行指令在被执行时使处理器执行如上的目标检测方法。

第四方面，本申请实施例还提供一种计算机可读存储介质，计算机可读存储介质存储一个或多个程序，一个或多个程序当被包括多个应用程序的电子设备执行时，使得电子设备执行如上的目标检测方法。

本申请实施例采用的上述至少一个技术方案能够达到以下有益效果：选择 YOLO-v4 构建目标检测模型，并基于待检测目标的特征设置针对降采样结构的调整方式，使得调整后得到的目标检测模型能够得到尺寸调整过的降采样特征图，在此基础上进行目标检测能够得到更高的精度。以工业上的缺陷检测场景为例，划痕、毛纤等成像为线状、体积较小的目标，如果利用原始降采样结构对检测图像进行处理，多次降采样会明显降低检测性能，而改良后的目标检测模型有效解决了这一问题。

附图简要说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

图 1 示出了 YOLO-v4 的网络结构示意图；

图 2 为在图 1 的网络结构基础上示出的各降采样层输出的特征图尺寸；

图 3 示出了根据本申请一个实施例的目标检测方法的流程示意图；

图 4 为在根据本申请一个实施例的目标检测模型的网络结构基础上示出的各降采样层输出的特征图尺寸；

图 5 为在根据本申请另一个实施例的目标检测模型的网络结构基础上示出的各降采样层输出的特征图尺寸；

图 6 为在根据本申请又一个实施例的目标检测模型的网络结构基础上示出的各降采样层输出的特征图尺寸；

图 7 为在根据本申请再一个实施例的目标检测模型的网络结构基础上示出的各降采样层输出的特征图尺寸；

图 8 示出了根据本申请一个实施例的目标检测模型的网络示意图；

图 9 示出了根据本申请另一个实施例的目标检测模型的网络示意图；

图 10 示出了根据本申请一个实施例的目标检测模型的结构示意图；

图 11 为本申请实施例中一种电子设备的结构示意图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

图 2 在图 1 示出的网络结构基础上示出了各降采样层输出的特征图尺寸。如图 2 所示，在输入图像的尺寸为 416×416 （单位为像素，下同），共分 RGB 三个通道（即图 2 中示出的 $416 \times 416 \times 3$ ，下面各层中标出的数字具有同类含义，不再一一解释）的情况下，先处理为 $416 \times 416 \times 32$ 的特征图（对应的网络结构在图 1 中未示出），然后依次经过 1/2 降采样层得到 208×208 的特征图、经过 1/4 降采样层得到 104×104 的特征图、经过 1/8 降采样层得到 52×52 的特征图、经过 1/16 降采样层得到 26×26 的特征图、经过 1/32 降采样层得到 13×13 的特征图。

原始 YOLO-v4 这样设计的原因在于，经过多次降采样可以得到尺寸较小的特征图，在较小的特征图上进行目标检测可以大大提高模型的推理速度。

但是发明人发现，这种方式对常见的成像为面状、且体积较大的自然物体

来说，多次降采样基本不会降低检测精度；但是，对于成像为线状、且体积较小的物体（尤其如工业检测中的一些细微毛纤缺陷，细微杂质缺陷）来说，多次降采样会明显降低模型的检测性能。

也就是说，现有技术的思路在于尽可能多的降采样，而本发明的设计思路在于，减少降采样，从而实现精度的提升。

以下结合附图，详细说明本申请各实施例提供的技术方案。

图 3 示出了根据本申请一个实施例的目标检测方法的流程示意图。如图 3 所示，该方法包括：

步骤 S310，基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式。

这里的待检测目标可以是车辆、缺陷等等各类需要检测的对象，而“特征”并不是指利用神经网络得到的张量特征，而是指细长、体积小等表象特征。为加以区别，后文中以“特征图”来表示利用神经网络得到的张量特征。

如前文所述，对于成像为线状、且体积较小的待检测目标，由于降采样次数过多会降低检测精度，因此此处的调整方式可以是减少降采样结构所实现的降采样次数或者是效果。

步骤 S320，利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整，构建基于 YOLO-v4 的目标检测模型。

步骤 S330，将检测图像输入到目标检测模型中，由目标检测模型提取检测图像的降采样特征图，根据降采样特征图得到目标检测结果；降采样特征图的尺寸是根据调整过的降采样结构确定的。

这里，对 YOLO-v4 的具体检测方式不做改变，可以参照图 1 和图 2，利用基于降采样层引出的检测分支，通过上采样、拼接、卷积等操作，利用锚框组，基于得到的降采样特征图得到目标检测结果。为了提升检测效果，也可以考虑增加检测分支，相应的实施例会在后面加以介绍。

可见，图 3 所示的方法，能够通过调整降采样结构来实现目标检测的精度提升，以工业上的缺陷检测场景为例，划痕、毛纤等成像为线状、体积较小的

目标，若利用原始降采样结构对检测图像进行处理，多次降采样会明显降低检测性能，而改良后的目标检测模型有效解决了这一问题。

在一些实施例中，目标检测方法中，利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整包括：对降采样结构中至少一个降采样层的步长进行调整。

例如，YOLO-v4 主干网络中，1/8 降采样层的步长（stride）为 2，参照图 2 可知，以检测图像的尺寸为 416*416 为例，通过 1/8 降采样层得到的降采样特征图的尺寸为 52*52。而如果将该步长由 2 调整为 1，则能够得到图 4 所示的效果，即以检测图像的尺寸为 416*416 为例，通过 1/8 降采样层得到的降采样特征图的尺寸仍为 104*104，与通过 1/4 降采样层得到的降采样特征图的尺寸相同（通道数发生变化）。

以原始 YOLO-v4 构建的目标检测模型作为对比例 1，以将 1/8 降采样层的步长设置为 1，其余不做变动得到的目标检测模型作为实施例 1，经过相同的样本集训练后，对实验集中的检测图像进行检测，实验数据表明实施例 1 在多个指标上均优于对比例 1，具体地，在平均检测精度（mAP）指标上具有约 6 个百分点的优势，在召回率（recall）指标上具有约 1 个百分点的优势，在检测精度（precision）指标上具有约 4 个百分点的优势。

在一些实施例中，目标检测方法中，利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整包括：删除降采样结构中的一个或多个降采样层。

删除降采样层，可以很直接地减少降采样次数。然而，减少降采样的次数也会使得降采样特征图的尺寸相对较大，这又会增加目标检测模型的训练和推理时间。

发明人通过实验，发现了较为均衡的方案。在一些实施例中，目标检测方法中，利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整包括：删除 1/4 降采样层和 1/32 降采样层中的任一个。图 5 示出了删除 1/4 降采样层后，各降采样层输出的特征图尺寸；图 6 示出了删除 1/32 降采样层后，各降采样层输出的特征图尺寸。

图 5 对应的目标检测模型有 255M 字节，图 6 对应的目标检测模型仅有 73.7M

字节，相较于前述对比例 1 的目标检测模型（256M 字节），体积也得到了减少，能够节约部署目标检测模型设备的内存空间。

在一些实施例中，目标检测方法中，利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整还包括：将原本连接在被删除的降采样层之后的各网络结构的通道数减半。例如，图 7 示出了删除 1/4 降采样层，并对其后各网络结构的通道数减半后，各降采样层输出的特征图尺寸。

图 7 对应的目标检测模型仅有 64.2M 字节，相较于前述对比例 1 的目标检测模型（256M 字节），体积也得到了减少，能够节约部署目标检测模型设备的内存空间。

在一些实施例中，目标检测方法中，构建基于 YOLO-v4 的目标检测模型包括：以调整后的降采样结构中的指定降采样层为基础，增加检测分支。

参照图 1 可知，原始 YOLO-v4 的主干网络具有三个检测分支，分别接在 1/8 降采样层、1/16 降采样层以及 1/32 降采样层后。

增加检测分支可以实现在更多的降采样特征图上进行目标检测，因此也可以提检测精度。例如，图 8 示出了根据本申请一个实施例的目标检测模型的网络示意图，相较于图 1 示出的原始 YOLO-v4 的主干网络，图 8 在 1/4 降采样层引出了一个新增的检测分支，即该目标检测模型共有四个检测分支，相较于原始 YOLO-v4 的主干网络，能够利用 1/4 降采样层得到的降采样特征图进行目标检测，实现精度的提升。

需要说明的是，增加检测分支的方案与前述减少降采样的方案可以结合使用，例如先对降采样结构进行调整，再基于调整后的降采样结构，调整检测分支的设置方式。如图 7 所示的网络结构，虽然删除了 1/4 降采样层，但是可以从 1/2 降采样层引出新增的检测分支；如图 5 所示的网络结构，虽然删除了 1/32 降采样层，但可以基于 1/2 降采样层、1/4 降采样层、1/8 降采样层和 1/16 降采样层引出检测分支，等等。

在一些实施例中，目标检测方法中，构建基于 YOLO-v4 的目标检测模型还包括：根据增加的检测分支，设置目标检测模型中各检测分支所使用的锚框。

锚框是在进行目标检测时选用的参考框，具体使用方式可以参照现有技术实现，本申请提出实施例的方案可以仅调整锚框的数量。

参照图 1 可知，原始 YOLO-v4 的主干网络中，由 1/8 降采样层引出的检测分支使用序号为 0、1、2 的三组锚框（anchor）组；由 1/16 降采样层引出的检测分支使用序号为 3、4、5 的三组锚框组；由 1/32 降采样层引出的检测分支使用序号为 6、7、8 的三组锚框组。

由于新增了检测分支，则需要确定新增的检测分支使用哪些锚框。

在一些实施例中，目标检测方法中，根据增加的检测分支，设置目标检测模型中各检测分支所使用的锚框包括：将第一预设数量的锚框组分配给各检测分支，第一预设数量是原始 YOLO-v4 主干网络所使用的锚框组数量，各检测分支被分配到的锚框组数量均不为 0；或者，将锚框组的数量由第一预设数量增加至第二预设数量，将第二预设数量的锚框组平均分配给各检测分支。

这里示出了两种可行的锚框分配方案，可以根据实际需求来选择其一来使用。在一种方案中，可以将原始 YOLO-v4 的主干网络中使用的 9 组锚框组重新分配给当前的所有检测分支，例如可以参照图 8，新增的检测分支（由 1/4 降采样层引出）使用序号为 0 的锚框组；由 1/8 降采样层引出的检测分支使用序号为 1、2、3 的三组锚框组；由 1/16 降采样层引出的检测分支使用序号为 4、5、6 的三组锚框组；由 1/32 降采样层引出的检测分支使用序号为 7、8 的两组锚框组。

或者，在另一种方案中，可以使各检测分支均使用相同数量的锚框组，例如参照图 9，示出了根据本申请另一个实施例的目标检测模型的网络示意图，在该实施例中，令新增的检测分支也使用三组锚框组，共使用 12 组锚框组。

在一些实施例中，目标检测方法还包括：对目标检测模型进行剪枝处理。

减少主干网络中的降采样和增加检测分支，虽然都可以提高对线状或者小体积目标的检测性能，但是也会带来更多的计算量。为了减少上述操作带来的计算量，以及降低网络过拟合的风险，可以对目标检测模型进行剪枝处理。

例如，可以选择 network slimming 剪枝算法，先对目标检测模型进行稀疏化训练，得到稀疏化的 γ 参数（前提是目标检测模型需要使用带 γ 参数的批标准

化 BN 层), 然后基于稀疏化的 γ 参数, 对卷积层的输入通道和/或输出通道进行剪枝处理。

本申请实施例还提供一种目标检测装置, 用于实现如上任一实施例提供的目标检测方法。具体地, 图 10 示出了根据本申请一个实施例的目标检测装置的结构示意图。如图 10 所示, 目标检测装置 1000 包括:

调整单元 1010, 用于基于待检测目标的特征, 对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式。

这里的待检测目标可以是车辆、缺陷等等各类需要检测的对象, 而“特征”并不是指利用神经网络得到的张量特征, 而是指细长、体积小等表象特征。为加以区别, 后文中以“特征图”来表示利用神经网络得到的张量特征。

如前文所述, 对于成像为线状、且体积较小的待检测目标, 由于降采样次数过多会降低检测精度, 因此此处的调整方式可以是减少降采样结构所实现的降采样次数或者是效果。

构建单元 1020, 用于利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整, 构建基于 YOLO-v4 的目标检测模型。

检测单元 1030, 用于将检测图像输入到目标检测模型中, 由目标检测模型提取检测图像的降采样特征图, 根据降采样特征图得到目标检测结果; 降采样特征图的尺寸是根据调整过的降采样结构确定的。

这里, 对 YOLO-v4 的具体检测方式不做改变, 可以参照图 1 和图 2, 利用基于降采样层引出的检测分支, 通过上采样、拼接、卷积等操作, 利用锚框组, 基于得到的降采样特征图得到目标检测结果。为了提升检测效果, 也可以考虑增加检测分支, 相应的实施例会在后面加以介绍。

可见, 图 10 所示的装置, 能够通过调整降采样结构来实现目标检测的精度提升, 以工业上的缺陷检测场景为例, 划痕、毛纤等成像为线状、体积较小的目标, 若利用原始降采样结构对检测图像进行处理, 多次降采样会明显降低检测性能, 而改良后的目标检测模型有效解决了这一问题。

在一些实施例中, 目标检测装置中, 构建单元 1020, 用于对降采样结构中

至少一个降采样层的步长进行调整。

在一些实施例中，目标检测装置中，构建单元 1020，用于删除降采样结构中的一个或多个降采样层。在一些实施例中，目标检测装置中，构建单元 1020，用于删除 $1/4$ 降采样层和 $1/32$ 降采样层中的任一个。在一些实施例中，目标检测装置中，构建单元 1020，用于将原本连接在被删除的降采样层之后的各网络结构的通道数减半。在一些实施例中，目标检测装置中，构建单元 1020，用于以调整后的降采样结构中的指定降采样层为基础，增加检测分支。在一些实施例中，目标检测装置中，构建单元 1020，用于根据增加的检测分支，设置目标检测模型中各检测分支所使用的锚框。在一些实施例中，目标检测装置中，构建单元 1020，用于将第一预设数量的锚框组分配给各检测分支，第一预设数量是原始 YOLO-v4 主干网络所使用的锚框组数量，各检测分支被分配到的锚框组数量均不为 0；或者，将锚框组的数量由第一预设数量增加至第二预设数量，将第二预设数量的锚框组平均分配给各检测分支。

在一些实施例中，目标检测装置还包括：剪枝单元，用于对目标检测模型进行剪枝处理。

能够理解，上述目标检测装置，能够实现前述实施例中提供的由目标检测服务器执行的目标检测方法的各个步骤，关于目标检测方法的相关阐释均适用于目标检测装置，此处不再赘述。

图 11 是本申请的一个实施例电子设备的结构示意图。请参考图 11，在硬件层面，该电子设备包括处理器，可选地还包括内部总线、网络接口、存储器。其中，存储器可能包含内存，例如高速随机存取存储器(Random-Access Memory, RAM)，也可能还包括非易失性存储器(non-volatile memory)，例如至少 1 个磁盘存储器等。当然，该电子设备还可能包括其他业务所需要的硬件。

处理器、网络接口和存储器可以通过内部总线相互连接，该内部总线可以是 ISA(Industry Standard Architecture, 工业标准体系结构)总线、PCI(Peripheral Component Interconnect, 外设部件互连标准)总线或 EISA(Extended Industry Standard Architecture, 扩展工业标准结构)总线等。总线可以分为地址总线、数

据总线、控制总线等。为便于表示，图 11 中仅用一个双向箭头表示，但并不表示仅有一根总线或一种类型的总线。

存储器，用于存放程序。具体地，程序可以包括程序代码，程序代码包括计算机操作指令。存储器可以包括内存和非易失性存储器，并向处理器提供指令和数据。

处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行，在逻辑层面上形成目标检测装置。图 11 示出的目标检测装置不构成本申请对目标检测装置数量的限制。处理器，执行存储器所存放的程序，并具体用于执行以下操作：

基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式；利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整，构建基于 YOLO-v4 的目标检测模型；将检测图像输入到目标检测模型中，由目标检测模型提取检测图像的降采样特征图，根据降采样特征图得到目标检测结果；降采样特征图的尺寸是根据调整过的降采样结构确定的。

上述如本申请图 1 所示实施例揭示的目标检测装置执行的方法可以应用于处理器中，或者由处理器实现。处理器可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器，包括中央处理器（Central Processing Unit, CPU）、网络处理器（Network Processor, NP）等；还可以是数字信号处理器（Digital Signal Processor, DSP）、专用集成电路（Application Specific Integrated Circuit, ASIC）、现场可编程门阵列（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。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。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。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电

可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成上述方法的步骤。

该电子设备还可执行图 1 中目标检测装置执行的方法，并实现目标检测装置在图 10 所示实施例的功能，本申请实施例在此不再赘述。

本申请实施例还提出了一种计算机可读存储介质，该计算机可读存储介质存储一个或多个程序，该一个或多个程序包括指令，该指令当被包括多个应用程序的电子设备执行时，能够使该电子设备执行图 1 所示实施例中目标检测装置执行的方法，并具体用于执行：

基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式；利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整，构建基于 YOLO-v4 的目标检测模型；将检测图像输入到目标检测模型中，由目标检测模型提取检测图像的降采样特征图，根据降采样特征图得到目标检测结果；降采样特征图的尺寸是根据调整过的降采样结构确定的。

本领域内的技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本申请是参照根据本申请实施例的方法、设备（系统）、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的

指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

在一个典型的配置中，计算设备包括一个或多个处理器 (CPU)、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器 (RAM) 和/或非易失性内存等形式，如只读存储器 (ROM) 或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。

以上所述仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求范围之内。

权利要求

1、一种目标检测方法，其特征在于，包括：

基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式；

利用所述调整方式对 YOLO-v4 主干网络的降采样结构进行调整，构建基于 YOLO-v4 的目标检测模型；

将检测图像输入到所述目标检测模型中，由所述目标检测模型提取所述检测图像的降采样特征图，根据所述降采样特征图得到目标检测结果；所述降采样特征图的尺寸是根据调整过的降采样结构确定的。

2、如权利要求 1 所述的方法，其特征在于，所述利用所述调整方式对 YOLO-v4 主干网络的降采样结构进行调整包括：

对所述降采样结构中至少一个降采样层的步长进行调整。

3、如权利要求 1 所述的方法，其特征在于，所述利用所述调整方式对 YOLO-v4 主干网络的降采样结构进行调整包括：

删除所述降采样结构中的一个或多个降采样层。

4、如权利要求 3 所述的方法，其特征在于，所述利用所述调整方式对 YOLO-v4 主干网络的降采样结构进行调整包括：

删除 1/4 降采样层和 1/32 降采样层中的任一个。

5、如权利要求 3 所述的方法，其特征在于，所述利用所述调整方式对 YOLO-v4 主干网络的降采样结构进行调整还包括：

将原本连接在被删除的降采样层之后的各网络结构的通道数减半。

6、如权利要求 1 所述的方法，其特征在于，所述构建基于 YOLO-v4 的目标检测模型包括：

以调整后的降采样结构中的指定降采样层为基础，增加检测分支。

7、如权利要求 6 所述的方法，其特征在于，所述构建基于 YOLO-v4 的目标检测模型还包括：

根据增加的检测分支，设置目标检测模型中各检测分支所使用的锚框。

8、如权利要求 7 所述的方法，其特征在于，所述根据增加的检测分支，设置目标检测模型中各检测分支所使用的锚框包括：

将第一预设数量的锚框组分配给各检测分支，所述第一预设数量是原始 YOLO-v4 主干网络所使用的锚框组数量，各检测分支被分配到的锚框组数量均不为 0；

或者，

将锚框组的数量由所述第一预设数量增加至第二预设数量，将第二预设数量的锚框组平均分配给各检测分支。

9、如权利要求 1~8 中任一项所述的方法，其特征在于，所述方法还包括：对所述目标检测模型进行剪枝处理。

10、一种目标检测装置，其特征在于，包括：

调整单元，用于基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式；

构建单元，用于利用所述调整方式对 YOLO-v4 主干网络的降采样结构进行调整，构建基于 YOLO-v4 的目标检测模型；

检测单元，用于将检测图像输入到所述目标检测模型中，由所述目标检测模型提取所述检测图像的降采样特征图，根据所述降采样特征图得到目标检测结果；所述降采样特征图的尺寸是根据调整过的降采样结构确定的。

11、如权利要求 10 所述的装置，其中，

所述构建单元，用于对降采样结构中至少一个降采样层的步长进行调整。

12、如权利要求 10 所述的装置，其中，

所述构建单元，用于删除降采样结构中的一个或多个降采样层。

13、如权利要求 12 所述的装置，其中，

所述构建单元，用于删除 1/4 降采样层和 1/32 降采样层中的任一个。

14、如权利要求 12 所述的装置，其中，

所述构建单元，用于将原本连接在被删除的降采样层之后的各网络结构的

通道数减半。

15、如权利要求 10 所述的装置，其中，

所述构建单元，用于以调整后的降采样结构中的指定降采样层为基础，增加检测分支。

16、一种电子设备，包括：处理器；以及被安排成存储计算机可执行指令的存储器，可执行指令在被执行时使处理器执行如下的目标检测方法：

基于待检测目标的特征，对 YOLO-v4 主干网络的降采样结构设置至少一个调整方式；利用调整方式对 YOLO-v4 主干网络的降采样结构进行调整，构建基于 YOLO-v4 的目标检测模型；将检测图像输入到目标检测模型中，由目标检测模型提取检测图像的降采样特征图，根据降采样特征图得到目标检测结果；降采样特征图的尺寸是根据调整过的降采样结构确定的。

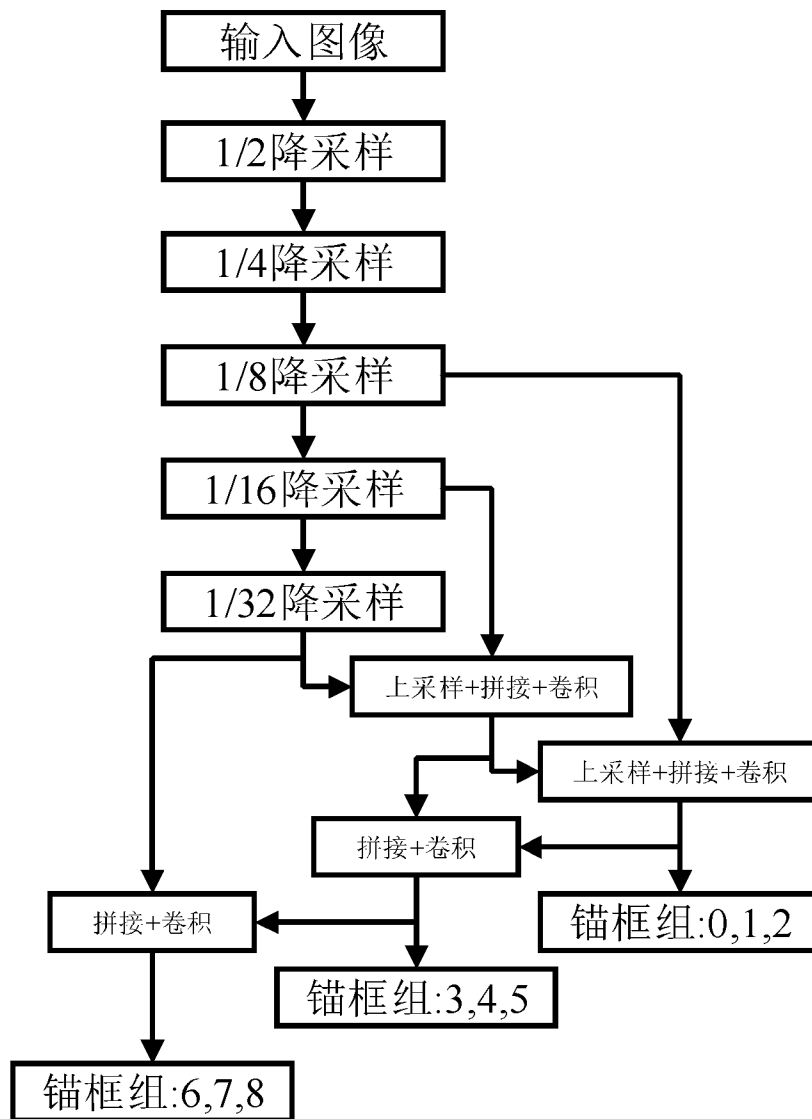


图 1

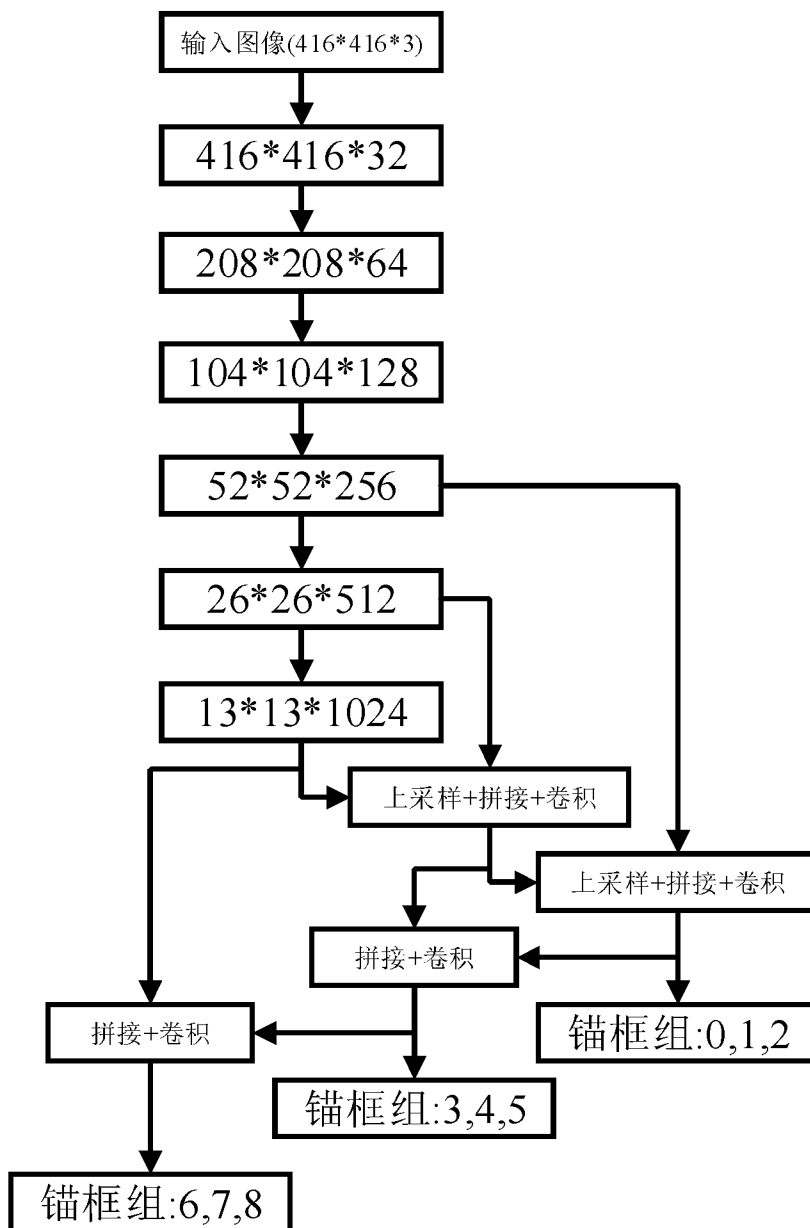


图 2

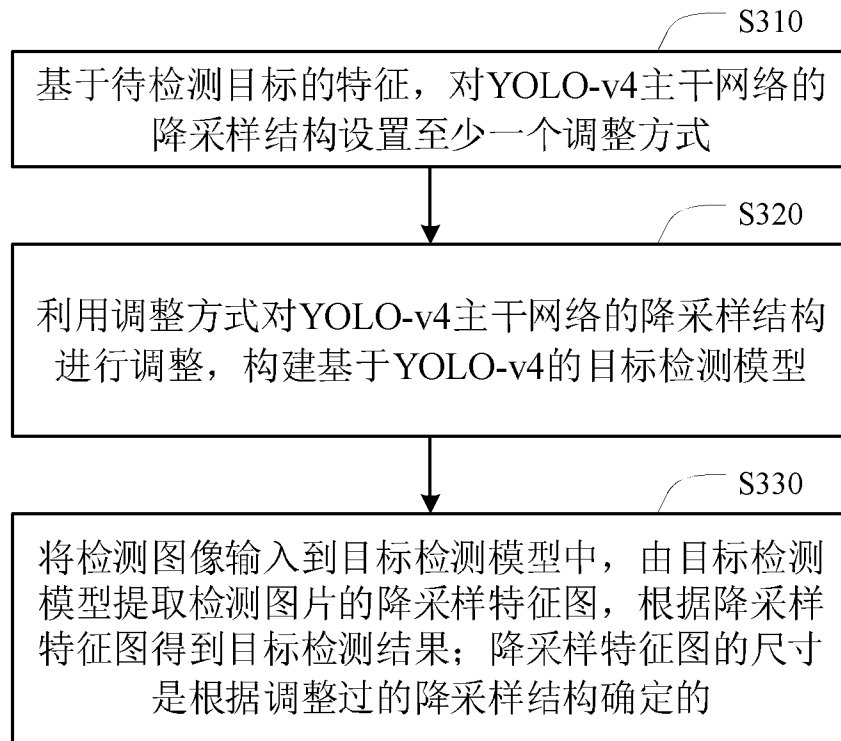


图 3

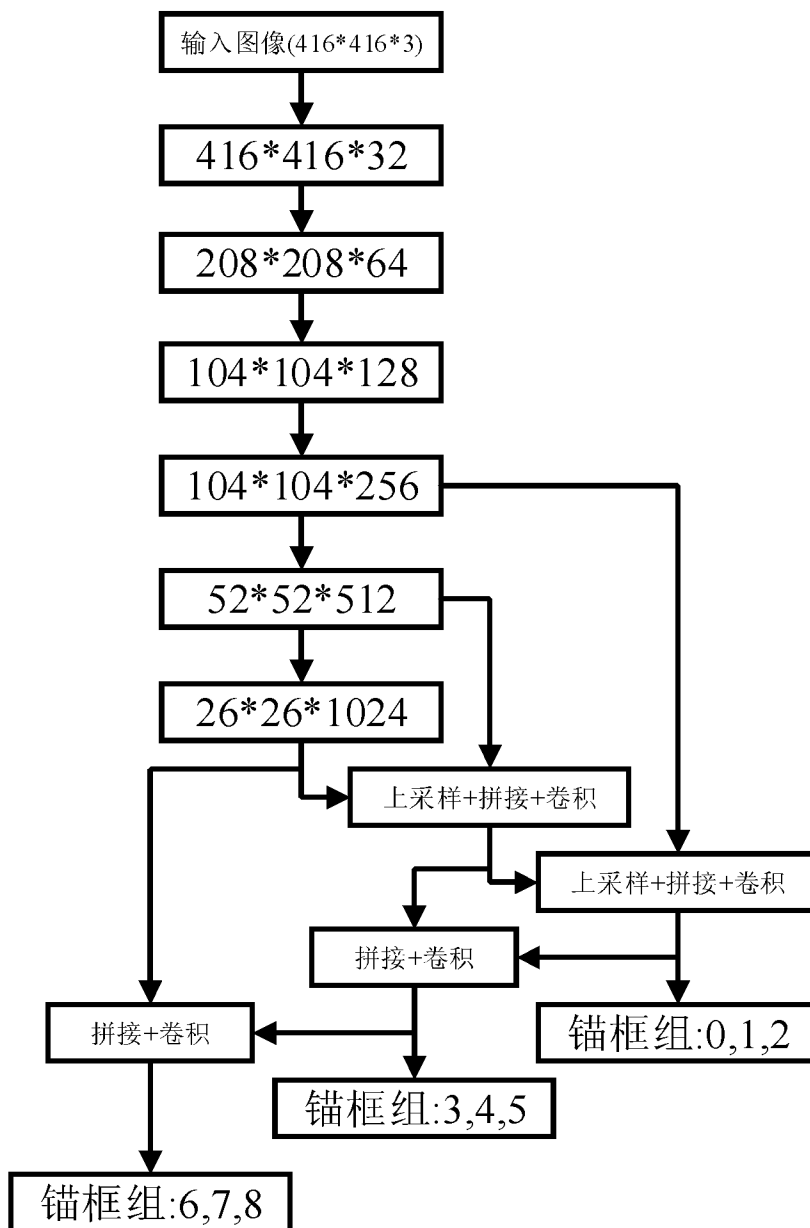


图 4

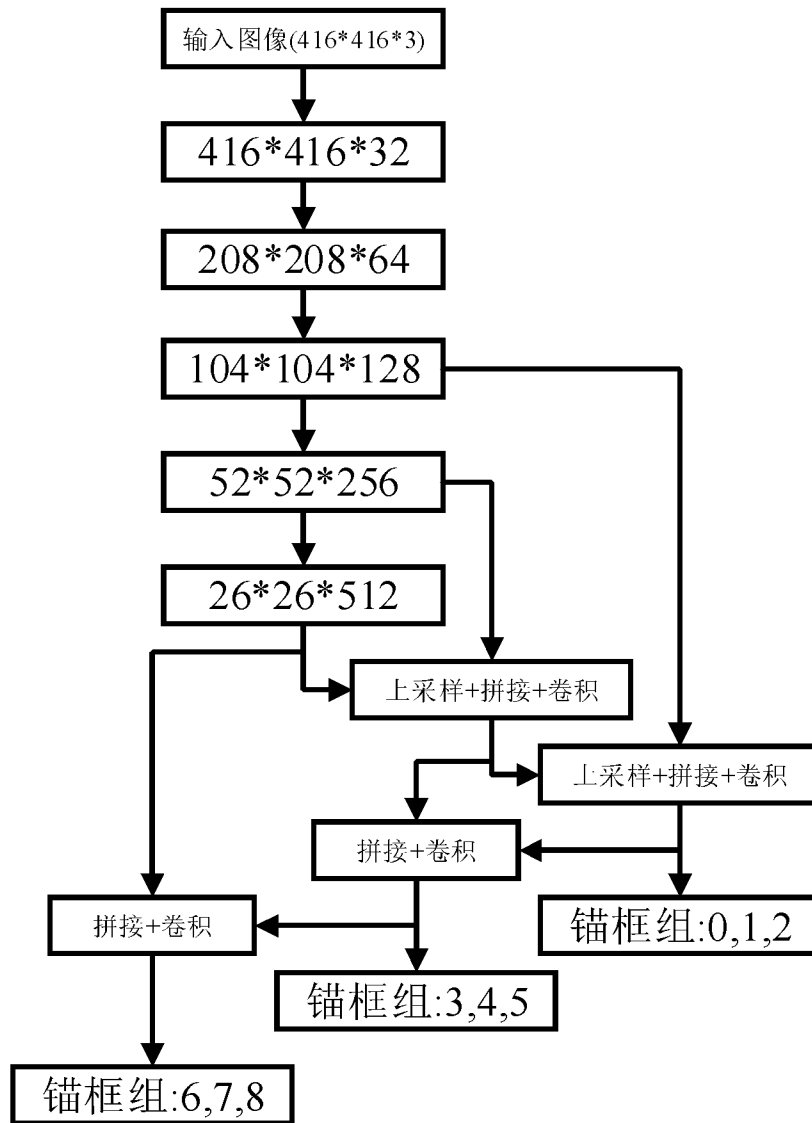


图 5

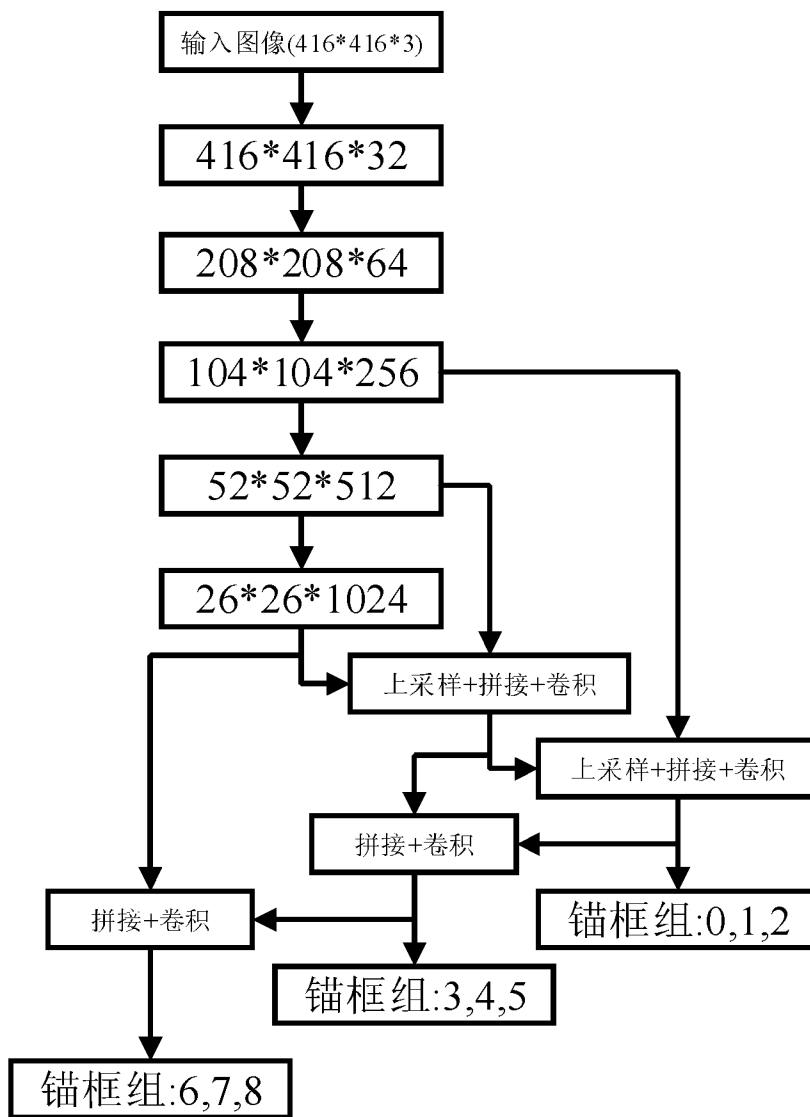


图 6

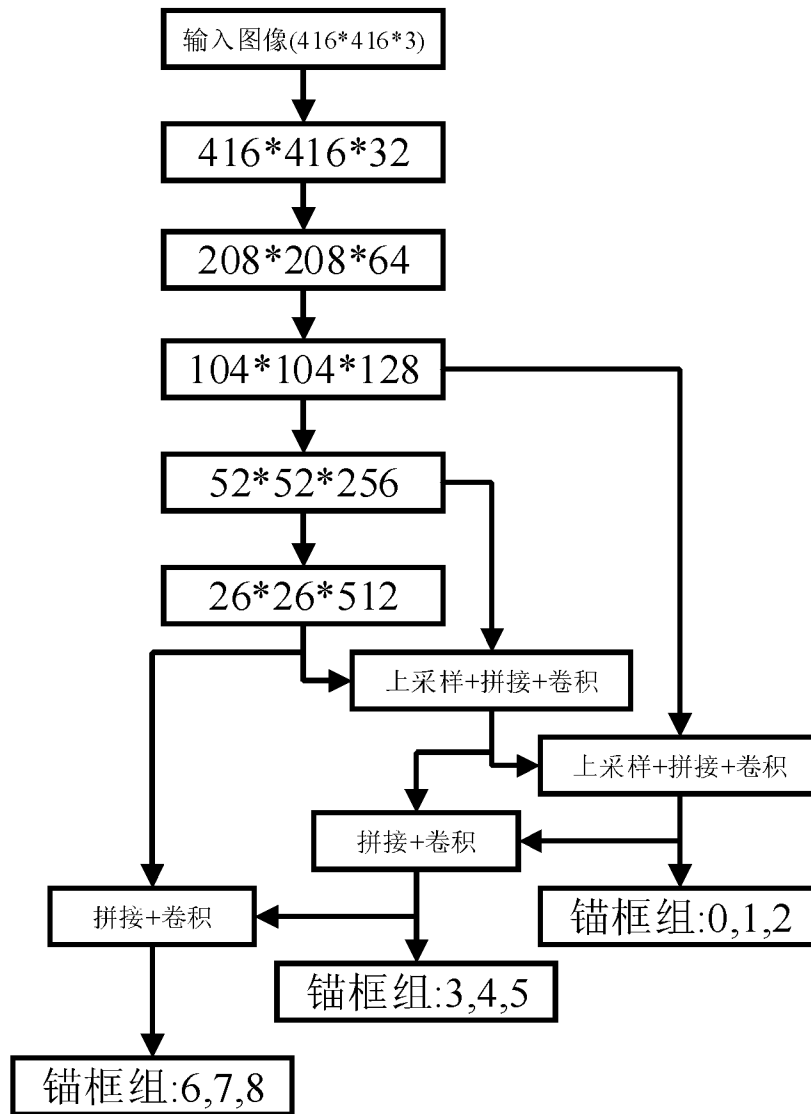


图 7

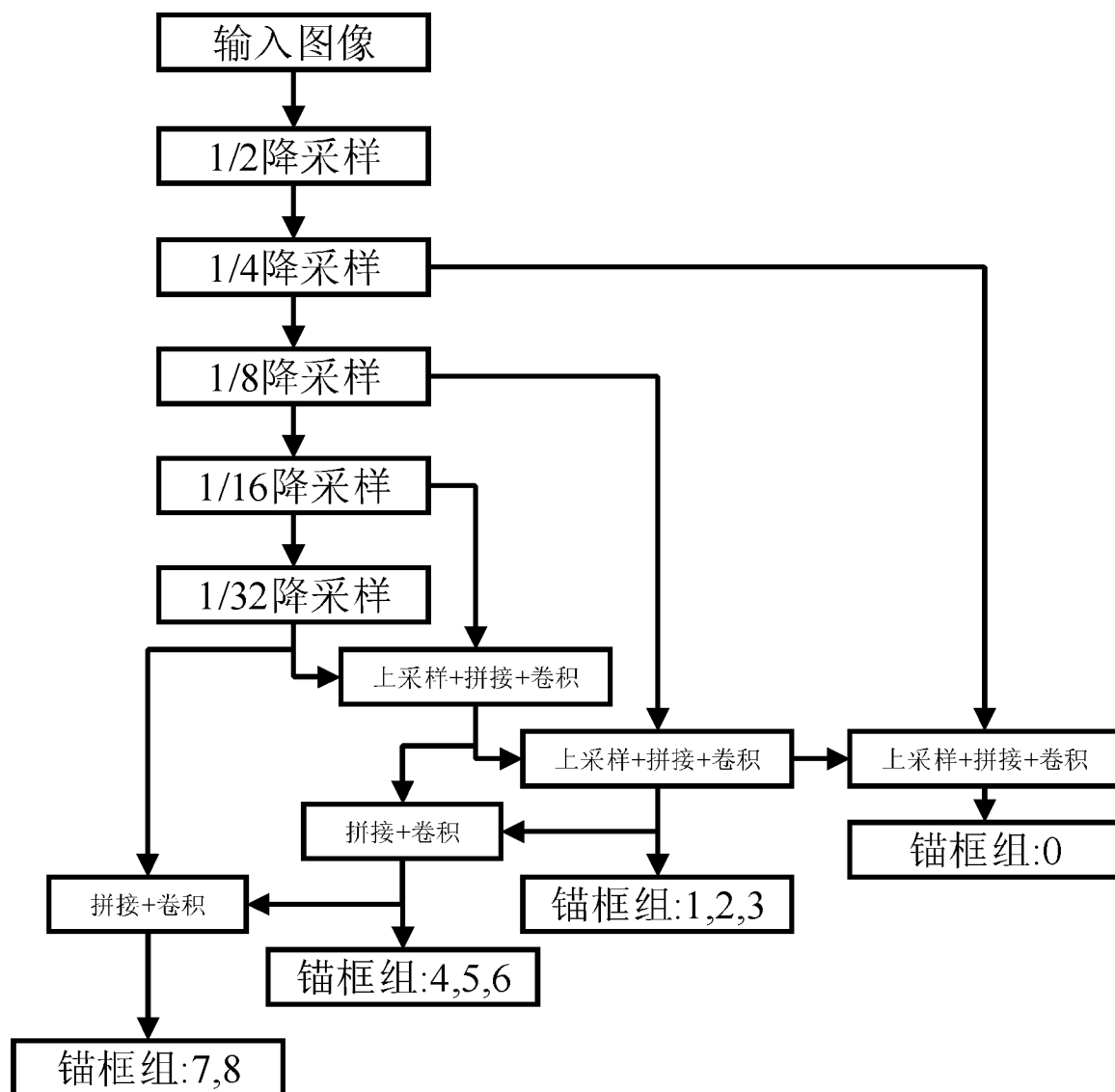


图 8

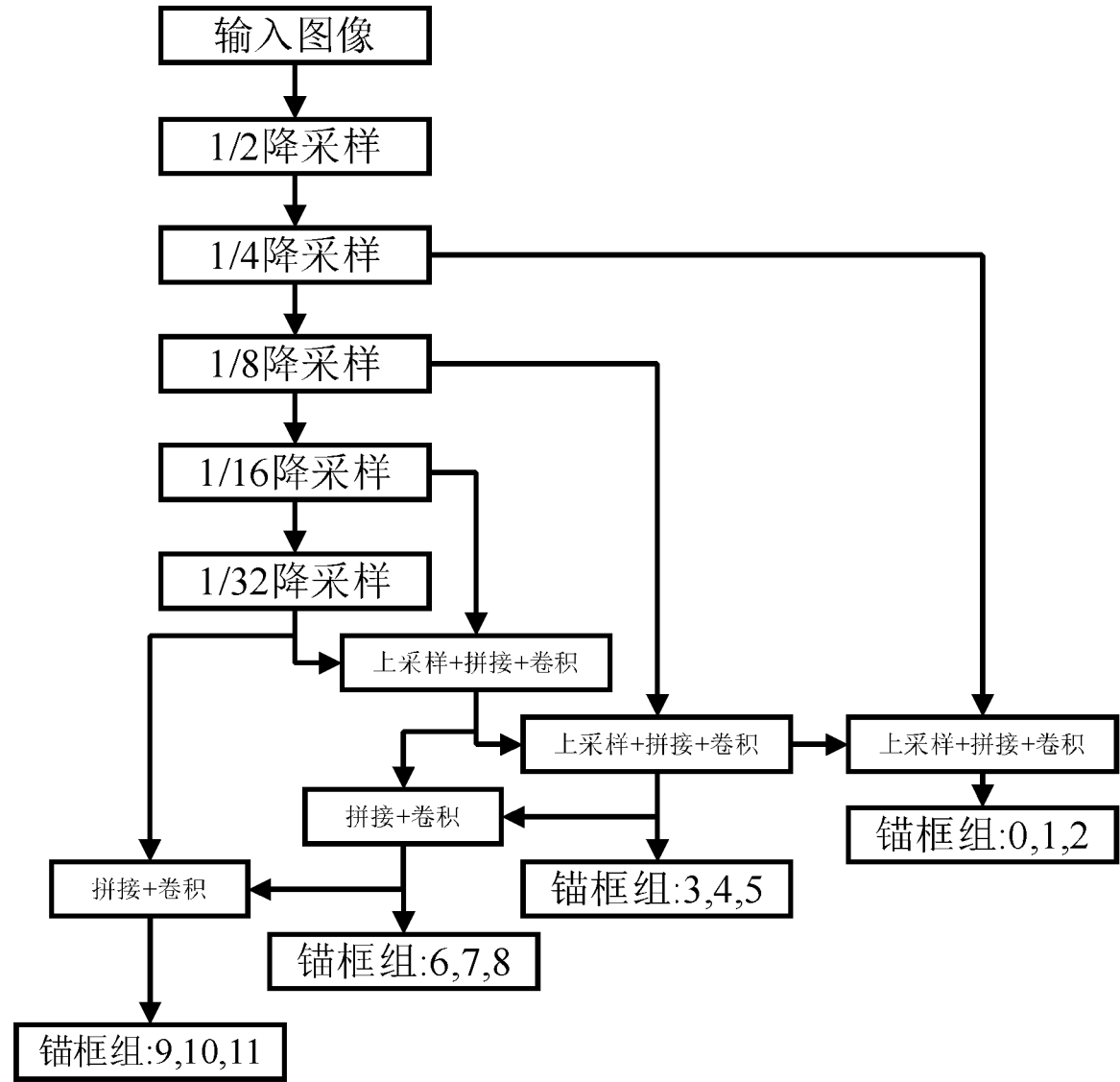


图 9

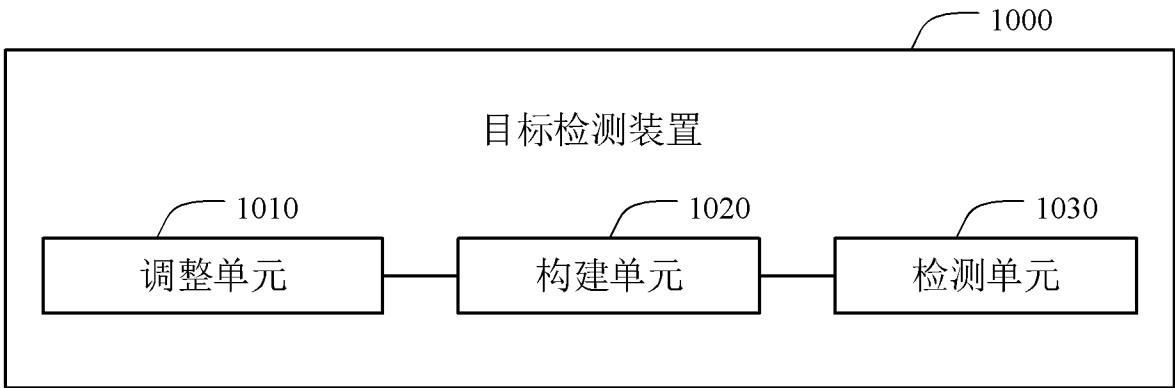


图 10

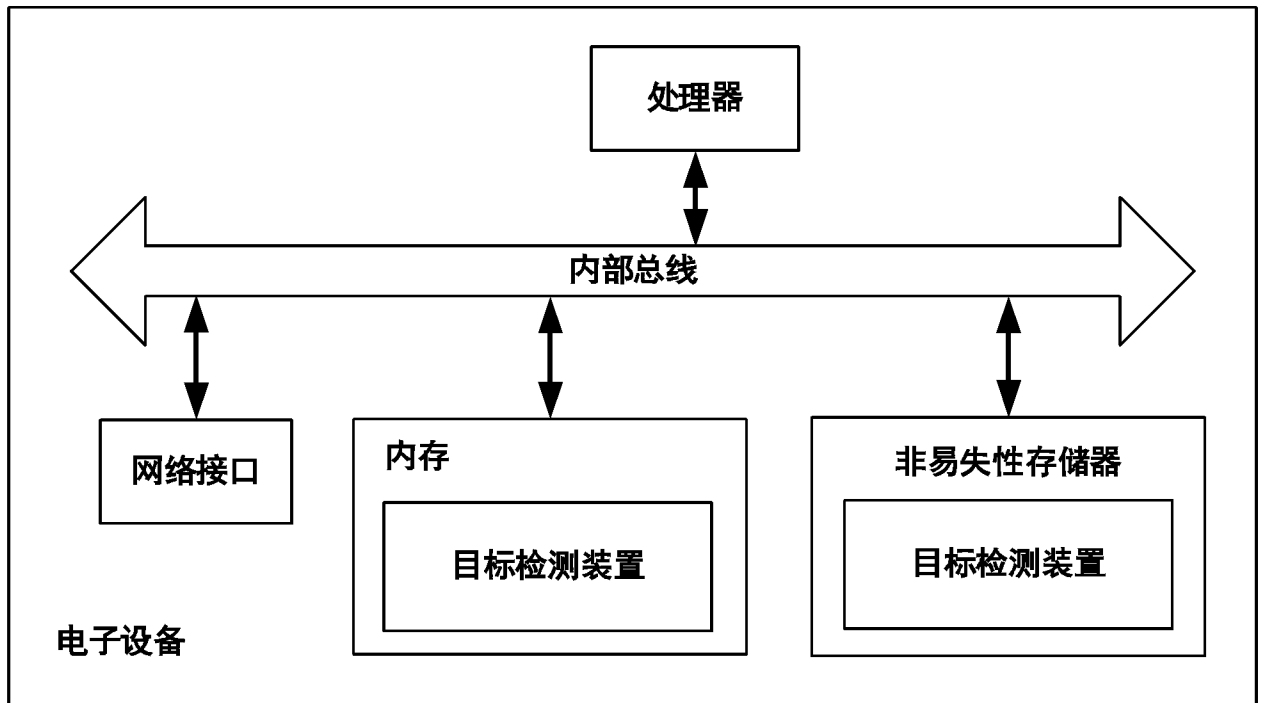


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/130102

A. CLASSIFICATION OF SUBJECT MATTER

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)

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)

CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT, CNKI: 目标检测, YOLO, 降采样, 特征, 尺寸, 层, 模型, 调整, target detection, down-sampling, characteristic, size, layer, model, adjust

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112949692 A (GOERTEK INC.) 11 June 2021 (2021-06-11) description, paragraphs 4-112, and figures 1-11	1-16
X	CN 111860064 A (HANGZHOU HIKVISION DIGITAL TECHNOLOGY CO., LTD.) 30 October 2020 (2020-10-30) description paragraphs 6-47, 57-117, figures 1-4	1-16
X	CN 111462050 A (UNIVERSITY OF SHANGHAI FOR SCIENCE AND TECHNOLOGY) 28 July 2020 (2020-07-28) description, paragraphs 27-62, figure 1	1-16
A	CN 110633594 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. et al.) 31 December 2019 (2019-12-31) entire document	1-16
A	US 2017206431 A1 (MICROSOFT TECHNOLOGY LICENSING L.L.C.) 20 July 2017 (2017-07-20) entire document	1-16



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

06 January 2022

Date of mailing of the international search report

28 January 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
Information on patent family members

International application No.

PCT/CN2021/130102

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112949692	A	11 June 2021	None			
CN	111860064	A	30 October 2020	None			
CN	111462050	A	28 July 2020	None			
CN	110633594	A	31 December 2019	None			
US	2017206431	A1	20 July 2017	US	9858496	B2	02 January 2018

国际检索报告

国际申请号

PCT/CN2021/130102

A. 主题的分类

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)

G06K

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, USTXT, WOTXT, EPTXT, CNKI: 目标检测, YOLO, 降采样, 特征, 尺寸, 层, 模型, 调整, target detection, down-sampling, characteristic, size, layer, model, adjust

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 112949692 A (歌尔股份有限公司) 2021年6月11日 (2021 - 06 - 11) 说明书第4-112段, 图1-11	1-16
X	CN 111860064 A (杭州海康威视数字技术股份有限公司) 2020年10月30日 (2020 - 10 - 30) 说明书第6-47、57-117段, 图1-4	1-16
X	CN 111462050 A (上海理工大学) 2020年7月28日 (2020 - 07 - 28) 说明书第27-62段, 图1	1-16
A	CN 110633594 A (北京京东尚科信息技术有限公司 等) 2019年12月31日 (2019 - 12 - 31) 全文	1-16
A	US 2017206431 A1 (MICROSOFT TECHNOLOGY LICENSING LLC) 2017年7月20日 (2017 - 07 - 20) 全文	1-16

☐ 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1月6日

国际检索报告邮寄日期

2022年1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

古志春

电话号码 86-(20)-28950386

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/130102

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	112949692	A	2021年6月11日	无	
CN	111860064	A	2020年10月30日	无	
CN	111462050	A	2020年7月28日	无	
CN	110633594	A	2019年12月31日	无	
US	2017206431	A1	2017年7月20日	US 9858496 B2	2018年1月2日