



(12) 发明专利申请

(10) 申请公布号 CN 113128362 A

(43) 申请公布日 2021. 07. 16

(21) 申请号 202110345378.5

(22) 申请日 2021.03.31

(71) 申请人 广东工业大学

地址 510062 广东省广州市越秀区东风东  
路729号

(72) 发明人 孟伟 胡扬 鲁仁全 麦达明

(74) 专利代理机构 广东广信君达律师事务所  
44329

代理人 张生梅

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

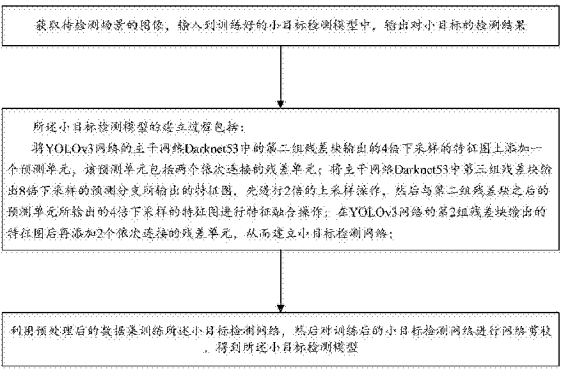
权利要求书2页 说明书7页 附图3页

(54) 发明名称

一种基于yolov3的无人机视角下小目标快速检测方法

(57) 摘要

本发明公开了一种基于yolov3的无人机视角下小目标快速检测方法,包括:获取待检测场景的图像,输入到训练好的小目标检测模型中,输出对小目标的检测结果;所述模型建立过程为:将Darknet53中的第二组残差块输出的4倍下采样的特征图上添加一个预测单元,该预测单元包括两个依次连接的残差单元;将第三组残差块输出8倍下采样的预测分支所输出的特征图,先进行2倍的上采样操作,然后与预测单元所输出的特征图进行特征融合操作;在YOLOv3网络的第2组残差块输出的特征图后再添加2个依次连接的残差单元,从而建立小目标检测网络;利用预处理后的数据集训练所述小目标检测网络,然后对训练后的小目标检测网络进行网络剪枝,得到所述小目标检测模型。



1. 一种基于yolov3的无人机视角下小目标快速检测方法,其特征在于,包括以下步骤:  
获取待检测场景的图像,输入到训练好的小目标检测模型中,输出对小目标的检测结果;

所述小目标检测模型的建立过程包括:

将YOLOv3网络的主干网络Darknet53中的第二组残差块输出的4倍下采样的特征图上添加一个预测单元,该预测单元包括两个依次连接的残差单元;将主干网络Darknet53中第三组残差块输出8倍下采样的预测分支所输出的特征图,先进行2倍的上采样操作,然后与第二组残差块之后的预测单元所输出的4倍下采样的特征图进行特征融合操作;在YOLOv3网络的第2组残差块输出的特征图后再添加2个依次连接的残差单元,从而建立小目标检测网络;

利用预处理后的数据集训练所述小目标检测网络,然后对训练后的小目标检测网络进行网络剪枝,得到所述小目标检测模型。

2. 根据权利要求1所述的基于yolov3的无人机视角下小目标快速检测方法,其特征在于,所述残差单元包括卷积核大小分别为 $1 \times 1$ 和 $3 \times 3$ 的两个卷积层。

3. 根据权利要求1所述的基于yolov3的无人机视角下小目标快速检测方法,其特征在于,所述数据集的预处理过程包括:

对于数据集中的图像,使用多次复制、粘贴小目标的方式以提高小目标样本的数量;

对数据集中的图像采用镜像反转,随机裁剪,随机旋转,随机缩放以及添加高斯噪声的方法进行数据增强。

4. 根据权利要求1所述的基于yolov3的无人机视角下小目标快速检测方法,其特征在于,对所述小目标检测网络进行训练时,所采用的损失函数如下:

$$L = \sum_{(x,y)} l(f(x,W),y) + \lambda \sum_{\gamma} g(\gamma)$$

其中第一项 $\sum_{(x,y)} l(f(x,W),y)$ 是网络预测所产生的各项损失, $x,y$ 表示的训练内的输入以及对应的target, $W$ 代表的是模型的训练权重;第二项 $g(\cdot)$ 代表的是对 $\gamma$ 因子的稀疏惩罚项, $\lambda$ 是权衡这两项损失的超参数。

5. 根据权利要求1所述的基于yolov3的无人机视角下小目标快速检测方法,其特征在于,在训练模型的过程中,首先将YOLOv3网络的主干网络在分类数据集Imagenet上预训练,使用交叉熵损失作为损失函数通过反向传播更新网络的权重;接着运用迁移学习的思想,将训练好的分类网络的部分权重作为小目标检测网络的主干网络的权重,然后数据集进行训练;损失函数包括两个部分,分别是分类损失和位置损失,对于分类损失采用交叉熵作为损失函数,位置损失则采用均方误差来作为损失函数;在训练的过程中,采用随机梯度下降方法来优化损失函数,采用余弦退火的方式来调整学习率,学习率初始化为0.001, Batch-size的大小是64,最大迭代次数epochs设置为10000。

6. 根据权利要求1所述的基于yolov3的无人机视角下小目标快速检测方法,其特征在于,采用训练后的小目标检测网络以及利用在卷积层后的批量归一化层中的 $\gamma$ 系数去衡量作为衡量卷积核重要性的标准,进而对卷积核进行裁剪;批量归一化层中的参数定义如下:

$$\hat{z} = \frac{z_{in} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

$$z_{out} = \gamma \hat{z} + \beta$$

其中,  $z_{in}$  代表前一层的输出,  $\mu_B$  代表前一层输出的均值,  $\sigma_B$  代表前一层输出的方差,  $\epsilon$  是趋于0的额外添加项; 将中心化的结果  $\hat{z}$  进行线性变换再作为下一层卷积层的输入  $z_{out}$ , 其中  $\gamma$  和  $\beta$  是用来施以线性变换的两个参数。

7. 根据权利要求1所述的基于yolov3的无人机视角下小目标快速检测方法, 其特征在于, 所述对训练后的小目标检测网络进行网络剪枝, 包括:

对训练后的小目标检测网络进行全局裁剪, 使用  $\gamma$  因子去衡量卷积核的重要性进而对全局卷积核中携带信息量较少的卷积核进行裁剪;

整个裁剪流程是基于  $\gamma$  大小实施的, 具体做法是:

首先对网络的全局  $\gamma$  的数值进行升序排序, 根据需要进行压缩参数的比例得到此次裁剪阈值  $T$ , 接着裁剪小于阈值  $T$  的  $\gamma$  因子所对应的卷积核, 设置一个单层的最大裁剪比例 0.5, 即单层卷积核的裁剪数量不超过50%; 通过引入L1正则化, 使得  $\gamma$  权重整体向零逼近;

接着是微调, 此时在数据集上进行恢复训练, 在保证精度的前提下将模型的参数进行裁剪;

重复裁剪和微调过程, 直至在精度与压缩比例达到设定的目标后停止。

8. 一种终端设备, 包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序, 处理器执行计算机程序时实现根据权利要求1至7中任一权利要求所述方法的步骤。

9. 一种计算机可读存储介质, 计算机可读存储介质存储有计算机程序, 其特征在于, 计算机程序被处理器执行时实现根据权利要求1至7中任一权利要求所述方法的步骤。

## 一种基于yolov3的无人机视角下小目标快速检测方法

### 技术领域

[0001] 本发明涉及目标检测技术领域,具体涉及一种基于yolov3的无人机视角下小目标快速检测方法,结合模型压缩技术,实现移动端嵌入式设备上的小尺度目标快速检测。

### 背景技术

[0002] 无人机目标检测作为一种新兴的技术,在航空图像分析、智能监控、路由检测等方面有着广泛的应用。目标检测近年来取得了很大的进展,特别是随着大规模视觉数据集的发展和计算能力的提高,深度神经网络(DNN),特别是卷积神经网络(CNN),在包括目标检测在内的计算机视觉任务中表现出了破记录的性能。然而,由于视角的特殊性目标尺度问题,它仍然是一项具有挑战性的工作。

[0003] 目前一些优秀的目标检测算法在大型通用的数据集上已经取得非常好的成绩,但是当前的目标检测算法基本上都是针对于自然场景中分布比较稀疏的中大型目标,所以导致对小目标的检测对目前来说仍是一个挑战。小目标由于分辨率较低,携带的信息十分有限,使得在经过多次下采样后小目标的分辨率进一步的降低,信息量少,所以通用的目标器具有对小目标检测精度低的特点。为了加强对小目标的检测,通常解决的一个办法就是利用特征金字塔的思想,将深层具有语义的特征和浅层的细节特征进行融合。浅层特征具有较小的感受野,对小目标的检测较好,而深层次的特征具有较大的感受野,对大尺度目标的检测较好,将不同层次的特征进行融合预测来解决对小目标检测的问题。

[0004] 其次就是虽然卷积神经网络相比传统的手工设计的特征算子提取的特征更加丰富,模型的泛化能力更强,但是神经网络往往是过参数化的,存在很多参数冗余,冗余的特征使得模型占用更多的资源,比如内存、CPU以及GPU等计算资源,所以在一些有限的硬件资源或对实时性要求较高的场景下,对模型进行压缩是不可避免的。一般模型压缩技术方法主要包括:低秩近似,网络剪枝,权值量化,知识蒸馏,设计更加紧凑的网络等技术。由于现在设计的神经网络中存在大量的尺寸为1x1大小的卷积核,极大限制了使用低秩近似的方法去压缩卷积核的参数量。

### 发明内容

[0005] 本发明的目的是提供一种基于yolov3的无人机视角下小目标快速检测方法,利用模型压缩技术,提高模型在端侧的检测速度的同时减少对小尺度目标的漏检率。

[0006] 为了实现上述任务,本发明采用以下技术方案:

[0007] 一种基于yolov3的无人机视角下小目标快速检测方法,包括以下步骤:

[0008] 获取待检测场景的图像,输入到训练好的小目标检测模型中,输出对小目标的检测结果;

[0009] 所述小目标检测模型的建立过程包括:

[0010] 将YOLOv3网络的主干网络Darknet53中的第二组残差块输出的4倍下采样的特征图上添加一个预测单元,该预测单元包括两个依次连接的残差单元;将主干网络Darknet53

中第三组残差块输出8倍下采样的预测分支所输出的特征图,先进行2倍的上采样操作,然后与第二组残差块之后的预测单元所输出的4倍下采样的特征图进行特征融合操作;在YOLOv3网络的第2组残差块输出的特征图后再添加2个依次连接的残差单元,从而建立小目标检测网络;

[0011] 利用预处理后的数据集训练所述小目标检测网络,然后对训练后的小目标检测网络进行网络剪枝,得到所述小目标检测模型。

[0012] 进一步地,所述残差单元包括卷积核大小分别为1\*1和3\*3的两个卷积层。

[0013] 进一步地,所述数据集的预处理过程包括:

[0014] 对于数据集中的图像,使用多次复制、粘贴小目标的方式以提高小目标样本的数量;

[0015] 对数据集中的图像采用镜像反转,随机裁剪,随机旋转,随机缩放以及添加高斯噪声的方法进行数据增强。

[0016] 进一步地,对所述小目标检测网络进行训练时,所采用的损失函数如下:

$$[0017] \quad L = \sum_{(x,y)} l(f(x, W), y) + \lambda \sum_{\gamma} g(\gamma)$$

[0018] 其中第一项  $\sum_{(x,y)} l(f(x, W), y)$  是网络预测所产生的各项损失,  $x, y$  表示的训练内的输入以及对应的target,  $W$  代表的是模型的训练权重;第二项  $g(\cdot)$  代表的是对  $\gamma$  因子的稀疏惩罚项,  $\lambda$  是权衡这两项损失的超参数。

[0019] 进一步地,在训练模型的过程中,首先将YOLOv3网络的主干网络在分类数据集Imagenet上预训练,使用交叉熵损失作为损失函数通过反向传播更新网络的权重;接着运用迁移学习的思想,将训练好的分类网络的部分权重作为小目标检测网络的主干网络的权重,然后数据集进行训练;损失函数包括两个部分,分别是分类损失和位置损失,对于分类损失采用交叉熵作为损失函数,位置损失则采用均方误差来作为损失函数;在训练的过程中,采用随机梯度下降方法来优化损失函数,采用余弦退火的方式来调整学习率,学习率初始化为0.001, Batch-size的大小是64,最大迭代次数epochs设置为10000。

[0020] 进一步地,采用训练后的小目标检测网络以及利用在卷积层后的批量归一化层中的  $\gamma$  系数去衡量作为衡量卷积核重要性的标准,进而对卷积核进行裁剪;批量归一化层中的参数定义如下:

$$[0021] \quad \hat{z} = \frac{z_{in} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

$$[0022] \quad z_{out} = \gamma \hat{z} + \beta$$

[0023] 其中,  $z_{in}$  代表前一层的输出,  $\mu_B$  代表前一层输出的均值,  $\sigma_B$  代表前一层输出的方差,  $\epsilon$  是趋于0的额外添加项;将中心化的结果  $\hat{z}$  进行线性变换再作为下一层卷积层的输入  $z_{out}$ , 其中  $\gamma$  和  $\beta$  是用来施以线性变换的两个参数。

[0024] 进一步地,所述对训练后的小目标检测网络进行网络剪枝,包括:

[0025] 对训练后的小目标检测网络进行全局裁剪,使用  $\gamma$  因子去衡量卷积核的重要性进而对全局卷积核中携带信息量较少的卷积核进行裁剪;

[0026] 整个裁剪流程是基于  $\gamma$  大小实施的,具体做法是:

[0027] 首先对网络的全局  $\gamma$  的数值进行升序排序,根据需要进行压缩参数的比例得到此次裁剪阈值T,接着裁剪小于阈值T的  $\gamma$  因子所对应的卷积核,设置一个单层的最大裁剪比例0.5,即单层卷积核的裁剪数量不超过50%;通过引入L1正则化,使得  $\gamma$  权重整体向零逼近;

[0028] 接着是微调,此时在数据集上进行恢复训练,在保证精度的前提下将模型的参数进行裁剪;

[0029] 重复裁剪和微调过程,直至在精度与压缩比例达到设定的目标后停止。

[0030] 一种终端设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,处理器执行计算机程序时实现前述基于yolov3的无人机视角下小目标快速检测方法的步骤。

[0031] 一种计算机可读存储介质,计算机可读存储介质存储有计算机程序,计算机程序被处理器执行时实现前述基于yolov3的无人机视角下小目标快速检测方法的步骤。

[0032] 与现有技术相比,本发明具有以下技术特点:

[0033] 1.本发明设计了一个针对无人机视角的小目标检测网络结构,提升对小尺度目标检测的鲁棒性;本发明实验中采用的是VisDrone数据集,该数据集在无人机视角下进行拍摄,包含10个目标类别,满足本次的应用场景需求。并且在数据预处理阶段采用复制粘贴的方式来增加小目标样本的数量,使网络的训练更集中于小目标。

[0034] 2.本发明在原有的YOLOv3算法基础上,引入更低的4倍下采样输出进行预测,增强小目标的检测;浅层的特征虽然语义程度较低,但是对位置的感知能力较好,并且浅层特征图的感受野较小,所以可以很好地用来对小目标检测。

[0035] 3.本发明结合模型压缩技术,实现在有限的硬件资源上,能够在移动端设备上对小目标进行快速的推理检测;通过引入L1正则化,使得  $\gamma$  权重整体向零逼近,进而实现对卷积核的大量裁剪,达到模型压缩的效果。

## 附图说明

[0036] 图1为本发明方法的流程示意图;

[0037] 图2为数据集中部分采集的图像以及无人机实验平台示意图;

[0038] 图3为通过复制粘贴增加小目标数量的示意图;

[0039] 图4的(a)是原有的YOLOv3网络结构图,(b)是改进后的YOLOv3网络结构图;

[0040] 图5是  $\gamma$  权重正则化的前后对比图;

[0041] 图6是本发明一个示例中的小目标检测结果示意图。

## 具体实施方式

[0042] 参见图1,本发明公开了一种基于yolov3的无人机视角下小目标快速检测方法,包括以下步骤:

[0043] 获取待检测场景的图像,输入到训练好的小目标检测模型中,输出对小目标的检测结果;

[0044] 所述小目标检测模型的建立过程包括:

[0045] 将YOLOv3网络的主干网络Darknet53中的第二组残差块输出的4倍下采样的特征图上添加一个预测单元,该预测单元包括两个依次连接的残差单元;将主干网络Darknet53中第三组残差块输出8倍下采样的预测分支所输出的特征图,先进行2倍的上采样操作,然后与第二组残差块之后的预测单元所输出的4倍下采样的特征图进行特征融合操作;在YOLOv3网络的第2组残差块输出的特征图后再添加2个依次连接的残差单元,从而建立小目标检测网络;

[0046] 利用预处理后的数据集训练所述小目标检测网络,然后对训练后的小目标检测网络进行网络剪枝,得到所述小目标检测模型。

[0047] 下面结合附图和具体示例对本发明的小目标检测模型作进一步详细说明。

[0048] 1.数据集的构建以及预处理

[0049] 本发明网络在训练时所用到的数据集中的部分图像如图2所示,本发明的实施例中使用装备分辨率为 $640 \times 480$ 的单目的无人机采集在不同场景下的图像以构建数据集,同时使用目标跟踪算法CSK在线自动标注数据集,尽可能的减少人力成本。

[0050] 针对自动标注的目标有限,将采集回来的数据集再次进行人工手动标注。此外,为了适应特殊的应用场景,将采集的数据分为训练集和验证集两个部分,将训练集用于模型的训练,验证集用于验证模型的检测的准确率。

[0051] 针对于数据集中的小目标样本数量较少的问题,通过在数据预处理时使用多次复制粘贴小目标来提高小目标样本的数量。此步骤旨在解决在网络训练过程中,由于小目标样本的数量较少,导致网络的注意力集中在中大型尺度的目标上,提升网络对小目标的拟合能力;也可以采用自适应重采样的数据扩充方式。此外为了增强模型的鲁棒性,所以在数据预处理阶段将训练的图像进行一系列的变换,包括镜像反转,随机裁剪,随机旋转,随机缩放以及添加高斯噪声等数据增强操作。

[0052] 2.小目标检测网络

[0053] 通常由于使用卷积神经网络对特征进行自动提取,由于感受野的问题,神经网络的浅层特征比较适合用于小尺度目标的检测但是语义信息较低,而较深的神经网络的特征具有较大的感受野并且语义信息较强,适合大尺度目标的检测,所以为了提升检测器对小目标的检测性能,因此本发明主要设计思想是通过浅层网络提取的特征的进一步复用,与深度网络提取的特征进行融合,从而增强对小目标检测的性能。为了实现深层次网络特征的复用,将高级其次采用多尺度预测并行预测的思想,实现在浅层网络的特征图上去预测小尺度目标,而在较深网络的特征图上去预测较大尺度的目标,通过这样的联合预测,让检测器能够对多尺度目标能有一个较好的检测性能。

[0054] 在本方案中的小目标检测网络是在YOLOv3网络的基础上改进得到的,YOLOv3主网络一共有23个残差块,共分为4组,每一组的残差块的个数分别是1、2、8、8、4,并且在每一组残差块后面都会进行2倍的下采样。为了更好的对小目标进行预测,对网络改进如下:

[0055] (1) 将YOLOv3网络的主干网络Darknet53中的第二组残差块输出的4倍下采样的特征图上添加一个预测单元,该预测单元是由两个依次连接的残差单元组成的,每个残差单元中是卷积核大小依次是 $1 \times 1$ 和 $3 \times 3$ 的两个卷积层。

[0056] (2) 将主干网络Darknet53中第三组残差块输出8倍下采样的预测分支所输出的特征图,先进行2倍的上采样操作,然后与第二组残差块后连接的预测分支输出的4倍下采样

的特征图进行特征融合操作,增强浅层特征图的语义信息。

[0057] (3) 为了获取更多低层的小目标位置信息,在YOLOv3网络的第2组残差块输出的特征图后再添加2个依次连接的残差单元。

[0058] 其中,第三组残差块固有四个预测分支,分别是4倍、8倍、16倍、32倍下采样的预测分支,该步骤中是对输出8倍下采样的预测分支所输出的特征图进行处理后,并与第二组残差块的预测分支共同建立输出为4倍降采样的特征融合目标层,以检测小目标。

[0059] 针对于改进后的网络,由于原YOLOv3中仅采用在8倍下采样的输出特征图的分支上进行预测,然后对小目标的检测效果并不是很好,而在本方案中通过添加在4倍下采样的预测分支,继而使得网络能够在训练过程中能够更好地对小目标进行拟合,由此使得在预测过程中减少对小目标预测的漏检率。

[0060] 同时,为了获取更多低层的小目标位置信息,在原网络的第2组残差块中增加2个残差单元,残差单元包括 $1 \times 1$ 卷积核和 $3 \times 3$ 卷积核, $1 \times 1$ 卷积核来对输入的通道数进行压缩,减少计算量,而 $3 \times 3$ 卷积核完成特征提取操作。采用该方法的优势是通过在更低倍下采样的特征图上进行预测,使得模型对小目标的检测能力进一步提升。

[0061] 3.网络训练过程

[0062] 由于在后续的模型压缩过程中采用的是基于Batch-Norm层中的 $\gamma$ 因子对卷积核进行裁剪,为了实现较大的压缩比,于是在原有的损失函数基础上增加一项关于 $\gamma$ 因子的L1正则项,公式定义如下所示:

$$[0063] \quad L = \sum_{(x,y)} l(f(x,W),y) + \lambda \sum_{\gamma} g(\gamma)$$

[0064] 其中第一项 $\sum_{(x,y)} l(f(x,W),y)$ 是模型预测所产生的各项损失, $x,y$ 表示的训练内的输入以及对应的target, $W$ 代表的是模型的训练权重。第二项 $g(\cdot)$ 代表的是对 $\gamma$ 因子的稀疏惩罚项, $\lambda$ 是权衡这两项损失的超参数,一般设置为 $1e-4$ 或者 $1e-5$ 。

[0065] 在训练模型的过程中,首先将YOLOv3网络的主干网络在分类数据集Imagenet上进行预训练,使用交叉熵损失作为损失函数通过反向传播更新网络的权重;接着运用迁移学习的思想,将训练好的分类网络的部分权重作为小目标检测网络的主干网络的权重,然后使用VisDrone无人机相关的数据集进行训练。目标检测需要对目标进行分类以及定位的操作,所以目标检测的损失函数也包括两个部分,分别是分类损失和位置损失,对于分类损失采用交叉熵作为损失函数,位置损失则采用均方误差来作为损失函数。在训练的过程中,采用随机梯度下降(SGD)方法来优化损失函数,采用余弦退火的方式来调整学习率,学习率初始化为0.001,Batch-size的大小是64,最大迭代次数epochs设置为10000。此次训练是在NVIDIA公司RTX 2080TI显卡上进行训练。

[0066] 4.网络剪枝

[0067] 在该步骤中,由于神经网络的庞大参数量以及内存高消耗的特点,所以势必要对模型进行优化才能进一步地在移动端嵌入式设备上部署。模型压缩的目的是为了最大程度的减小模型的复杂度,减少模型存储所需要的空间以及在推理时对内存的消耗,加速模型的训练和推理过程。本发明采用基于现有的预训练模型以及利用常在卷积层后的批量归一化层(BN-layer)中的 $\gamma$ 系数去衡量作为衡量卷积核重要性的标准,进而对卷积核进行



裁剪。批量归一化层中的参数定义如下：

$$[0068] \quad \hat{z} = \frac{z_{in} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

$$[0069] \quad z_{out} = \gamma \hat{z} + \beta$$

[0070] 其中,  $z_{in}$  代表前一层的输出,  $\mu_B$  代表前一层输出的均值,  $\sigma_B$  代表前一层输出的方差,  $\epsilon$  是趋于0的额外添加项; 将中心化的结果  $\hat{z}$  再进行线性变换再作为下一层卷积层的输入  $z_{out}$ , 其中  $\gamma$  和  $\beta$  是用来施以线性变换的两个参数, 目的是让神经网络自己去学习这两个参数从而解决数据分布差异较大的问题。本发明中采用的是基于批量归一化层  $\gamma$  因子来对卷积核进行裁剪, 该裁剪方式属于结构化裁剪, 相比于非结构化裁剪, 这种方式能够在现有的模型训练框架下进行快速的推理, 不需要特殊的硬件或软件, 具体过程如下:

[0071] 对训练后的小目标检测网络进行全局裁剪, 使用  $\gamma$  因子去衡量卷积核的重要性进而对全局卷积核中携带信息量较少的卷积核进行裁剪, 由于对  $\gamma$  引入了L1正则化约束, 所以大量的  $\gamma$  权重会逼近0。整个裁剪流程是基于  $\gamma$  大小实施的, 具体做法是:

[0072] 首先对网络的全局  $\gamma$  的数值进行升序排序, 根据需要进行压缩参数的比例得到此次裁剪阈值T, 接着裁剪小于阈值T的  $\gamma$  因子所对应的卷积核; 考虑到单个卷积层中的卷积核数量可能会出现全被裁减掉的情况, 导致在后续的微调阶段也无法恢复模型的精度, 所以在实验中设置一个单层的最大裁剪比例0.5, 即单层卷积核的裁剪数量不超过50%。

[0073] 由于将网络的参数进行了裁剪, 使得模型的检测精度会下降, 此时在训练集上进行恢复训练, 在保证精度的前提下将模型的参数进行裁剪。

[0074] 最后重复裁剪和微调过程, 直至在精度与压缩比例有一个好的折中后停止训练; 除了上述此次用到的模型结构化裁剪技术, 现有的模型压缩技术还包括权重的量化以及知识蒸馏等技术。网络剪枝又包括结构化剪枝和非结构化的剪枝。对于非结构化剪枝就是针对权重直接进行裁剪, 将值比较小的权重置零, 这样就可以得到网络的系稀疏结构, 但是这种结构需要有专门的软件或者硬件才能有很大的加速效果。

[0075] 本发明通过使用新的数据增强技术, 来进一步提升网络对小目标的拟合能力。

[0076] 通常在一些大型通用的数据集中(例如COCO), 小目标样本的数量较少, 使得只有在IOU阈值设置较低的情况下才能匹配到Anchor。并且在训练的过程中, 由于小目标的损失对整体损失的贡献较少, 使得网络的注意力集中在中大型的目标拟合上, 导致小目标检测精度较低。如图3所示, 在包含小目标的样本中多次复制粘贴小目标来解决。当粘贴每个目标的时候, 需要保证粘贴的目标与现有的目标不重叠, 增加了小目标位置的多样性, 同时保证目标出现在合适的上下文, 每个样本中小目标个数的增加进一步解决了匹配的anchor数目问题。通过数据增强技术, 实现在检测时能够对目标出现遮挡, 不同尺度大小等情况有一个良好的鲁棒性。

[0077] 本发明融合更浅层特征加强了对小目标的预测能力。

[0078] 如图4所示, (a) 是原有的网络结构, (b) 是改进后的网络; YOLOv3的网络采用在8倍的下采样的特征图上进行预测的方式来对小目标进行检测, 由于浅层特征包含更多小目标的位置信息, 但语义特征较低, 所以为了进一步的加强对小目标的预测能力, 在原来的基础之上输出一个4倍下采样的特征图, 并且将具有高级语义信息的特征图进行特征融合, 实现在无人机视角增强对小尺度目标的检测性能, 也包括对不同类别的目标的检测, 模型有较

好的泛化能力;在一个示例中,检测结果如图6所示。

[0079] 通过使用模型剪枝的技术,提升速模在端侧的检测速度,实测推理速度为13FPS/s,达到实时性检测要求;YOLOv3网络具有53个卷积层并且采用了多尺度的预测方式,浮点运算量以及内存消耗都非常大,导致模型在移动端的推理速度非常慢,在Nvidia TX2嵌入式设备上实测2~3FPS/s,所以无法满足实时检测的应用需求。本方案首先使用 $\gamma$ 权重来衡量卷积核的重要性,并对卷积核的裁剪,使得模型在保持精度的前提下大量压缩模型参数。如图5所示, $\gamma$ 参数在未压缩之前,总体上近似正态分布,完成 $\gamma$ 对卷积核的选择人物任务,所以通过对 $\gamma$ 因子引入L1正则化以后, $\gamma$ 参数整体向0逼近,通过设置的阈值可以将大量的卷积核进行裁剪,大大的减少模型的参数量,检测速度大大提升,在Nvidia TX2嵌入式设备上实测16FPS/s,满足应用要求。如表1所示,压缩后的YOLOv3在检测精度上虽然略有降低但是比YOLOv3的Tiny版本的检测精度要高,并且在检测速度上有更大的优势。

[0080] 表1算法性能对比

[0081]

Metric	YOLOv3	YOLOv3-Tiny	改进后的YOLOv3
mAP	0.297	0.136	0.167
FPS	2	7	16
模型大小	237M	31M	12.2M

[0082] 本申请实施例进一步提供一种终端设备,该终端设备可以为计算机、服务器;包括存储器、处理器以及存储在存储器中并可在处理器上运行的计算机程序,处理器执行计算机程序时实现上述基于yolov3的无人机视角下小目标快速检测方法的步骤。

[0083] 本申请的实施提供了一种计算机可读存储介质,计算机可读存储介质存储有计算机程序,计算机程序被处理器执行时实现上述基于yolov3的无人机视角下小目标快速检测方法的步骤。

[0084] 计算机程序也可以被分割成一个或多个模块/单元,一个或者多个模块/单元被存储在存储器中,并由处理器执行,以完成本申请。一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段,该指令段用于描述计算机程序在终端设备中的执行过程。

[0085] 以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围,均应包含在本申请的保护范围之内。

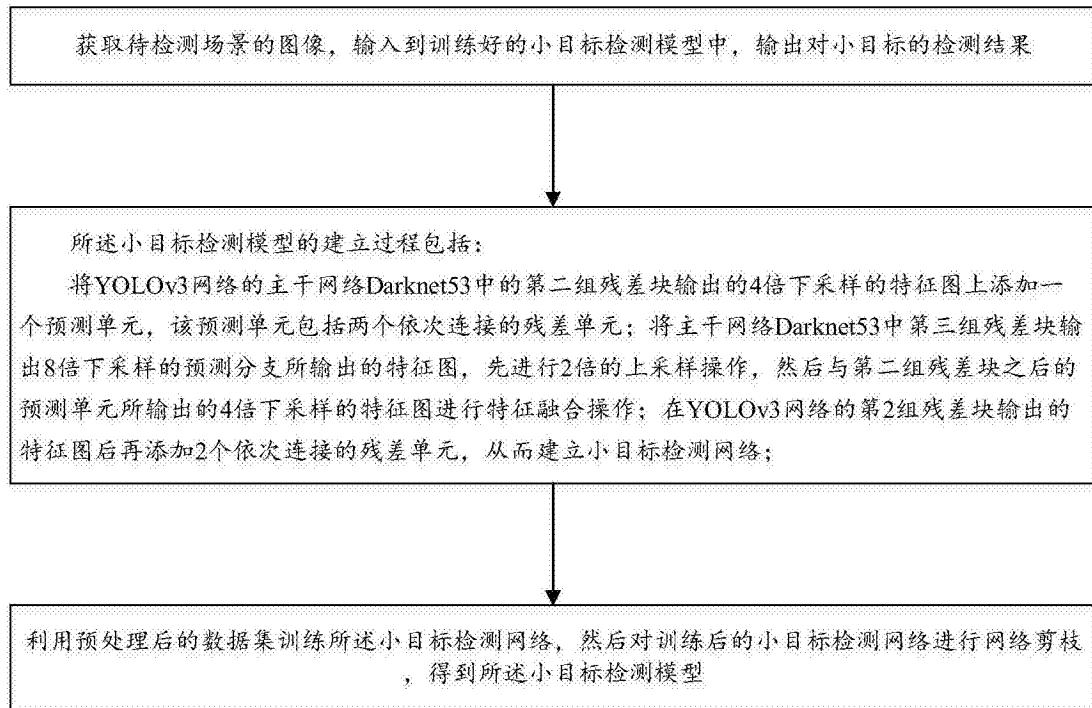


图1



图2

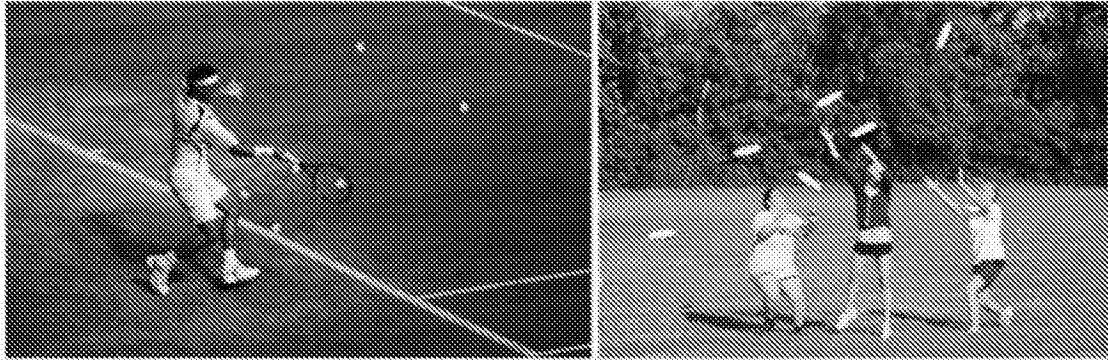


图3

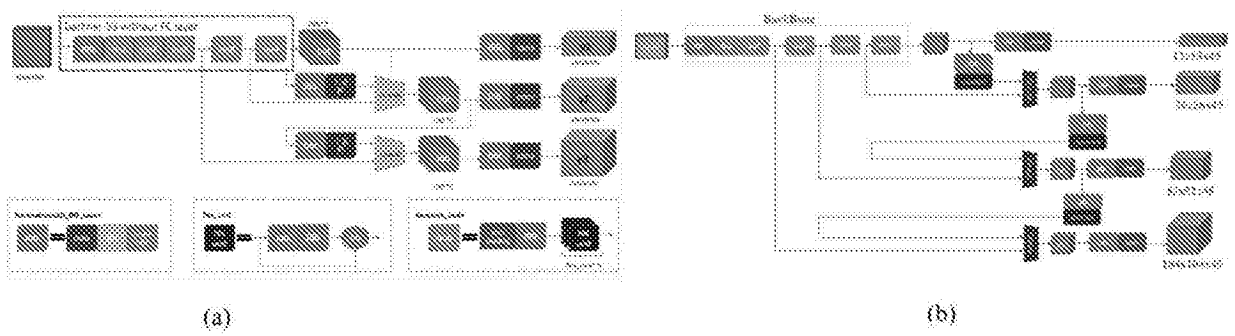


图4

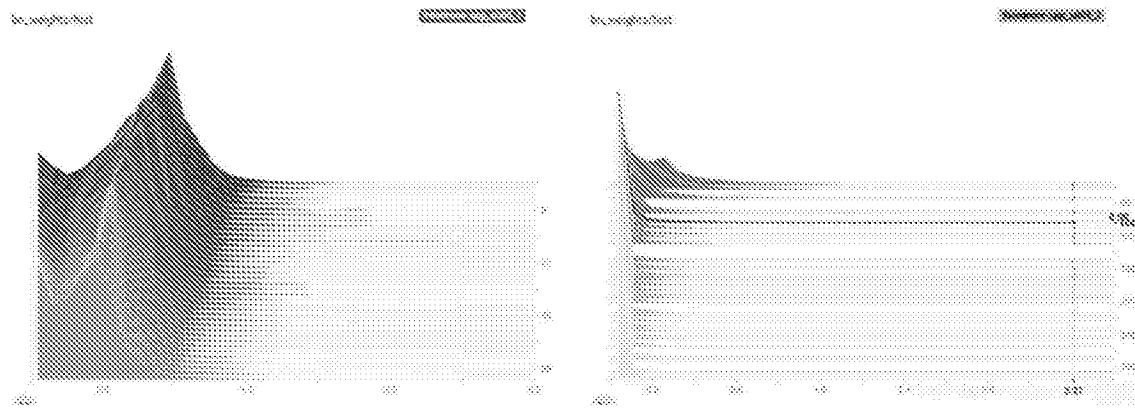


图5



图6