



(12) 发明专利申请

(10) 申请公布号 CN 112364843 A

(43) 申请公布日 2021.02.12

(21) 申请号 202110029055.5

G06T 7/70 (2017.01)

(22) 申请日 2021.01.11

(71) 申请人 中国科学院自动化研究所

地址 100190 北京市海淀区中关村东路95号

(72) 发明人 兰晓松 李书晓 朱承飞

(74) 专利代理机构 北京市恒有知识产权代理事

务所(普通合伙) 11576

代理人 郭文浩 尹文会

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/32 (2006.01)

G06K 9/62 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

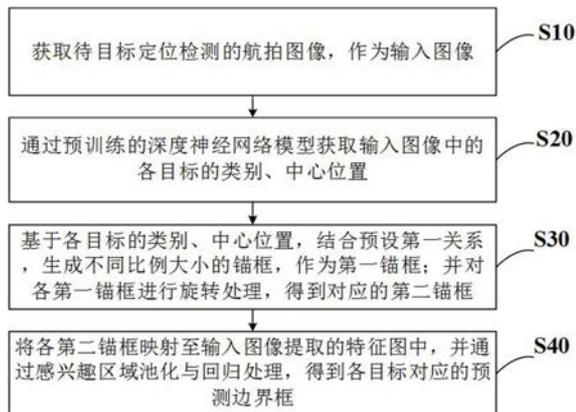
权利要求书2页 说明书9页 附图5页

(54) 发明名称

插拔式航拍图像目标定位检测方法、系统、设备

(57) 摘要

本发明属于计算机视觉技术领域,具体涉及一种插拔式航拍图像目标定位检测方法、系统、设备,旨在解决航拍图像目标定位检测精度、识别精度及效率较低的问题。本方法包括获取待定位检测的航拍图像,作为输入图像;通过预训练的神经网络模型获取输入图像中的各目标的类别、中心位置;基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;将各第二锚框映射至输入图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。本发明提高了航拍图像目标定位检测精度、识别精度及效率。



1. 一种插拔式航拍图像目标定位检测方法,其特征在于,该方法包括以下步骤:

步骤S10,获取待定位检测的航拍图像,作为输入图像;

步骤S20,通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置;

步骤S30,基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

步骤S40,将各第二锚框映射至所述输入图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

2. 根据权利要求1所述的插拔式航拍图像目标定位检测方法,其特征在于,所述神经网络模型基于残差模块、第一卷积分支、第二卷积分支构建;所述第一卷积分支和第二卷积分支均包括一个 3×3 卷积层、一个 1×1 卷积层;其中,所述第一卷积分支用于生成目标中心点的热力图;所述第二个卷积分支用于生成目标中心点的偏移图。

3. 根据权利要求2所述的插拔式航拍图像目标定位检测方法,其特征在于,“通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置”,其方法为:

提取所述输入图像的特征,并分别通过第一卷积分支、第二卷积分支进行卷积处理,得到各目标的目标中心点的热力图、偏移图;

基于各目标的目标中心点的热力图,通过非极大值抑制方法NMS得到各目标的类别、中心位置;并根据所述目标中心点的偏移图对中心位置进行调整。

4. 根据权利要求3所述的插拔式航拍图像目标定位检测方法,其特征在于,所述神经网络模型其在训练过程中的损失函数为:

$$L = L_k + L_{off}$$

$$L_k = \frac{-1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}) & \text{if } Y_{xyc} = 1 \\ (1 - \hat{Y}_{xyc})^\beta (\hat{Y}_{xyc})^\alpha & \text{otherwise} \\ \log(1 - \hat{Y}_{xyc}) & \end{cases}$$

$$L_{off} = \frac{1}{N} \sum_p \left| \hat{o}_{\tilde{p}} - \left(\frac{p}{R} - \tilde{p} \right) \right|$$

其中, L 表示总损失, L_k 表示目标中心热力图损失, L_{off} 表示目标中心偏移损失, N 为目标中心点个数, α 、 β 为Focal Loss的超参, \hat{Y}_{xyc} 表示检测到的目标中心点置信度, Y_{xyc} 表示高斯核函数, p 表示每个目标类别的所有目标中心点, R 为输出步长, $\tilde{p} = \lfloor \frac{p}{R} \rfloor$, \hat{o} 为预测的中心点局部偏移。

5. 根据权利要求1所述的插拔式航拍图像目标定位检测方法,其特征在于,步骤S30中“对各第一锚框进行旋转处理”,其方法为:以各第一锚框的框中心为旋转中心,长边作为X

轴,短边作为Y轴,在0-180°范围内每M度旋转一次;M为每一次旋转的角度,为正整数。

6. 一种插拔式航拍图像目标定位检测系统,其特征在于,该系统包括:图像获取模块、类别及位置获取模块、锚框预处理模块、边界框预测模块;

所述图像获取模块,配置为获取待定位检测的航拍图像,作为输入图像;

所述类别及位置获取模块,配置为通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置;

所述锚框预处理模块,配置为基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

所述边界框预测模块,配置为将各第二锚框映射至所述输入图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

7. 一种航拍图像目标定位检测系统,其特征在于,该系统包括:提取模块、锚框预处理模块、边界框预测模块;

所述提取模块,配置为获取待目标定位检测的航拍图像,并提取该航拍图像中各目标的类别、中心位置;

所述锚框预处理模块,配置为基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

所述边界框预测模块,配置为将各第二锚框映射至所述航拍图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

8. 一种插拔式航拍图像目标定位检测设备,其特征在于,包括:

至少一个处理器;以及

与至少一个所述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处理器执行的指令,所述指令用于被所述处理器执行以实现权利要求1-5任一项所述的插拔式航拍图像目标定位检测方法。

9. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有计算机指令,所述计算机指令用于被所述计算机执行以实现权利要求1-5任一项所述的插拔式航拍图像目标定位检测方法。

插拔式航拍图像目标定位检测方法、系统、设备

技术领域

[0001] 本发明属于计算机视觉技术领域,具体涉及一种插拔式航拍图像目标定位检测方法、系统、设备。

背景技术

[0002] 目前主流的目标检测算法主要是基于深度学习模型,大多是针对通用物体的水平检测识别,大概可以分成两大类:(1)One-Stage目标检测算法,这类检测算法不需要Region Proposal阶段,可以通过一个阶段直接产生物体的类别概率和位置坐标值,比较典型的算法有YOLO、SSD、RetinaNet、CornerNet和CenterNet等(参见文献:“Zou Z, Shi Z, Guo Y, et al. Object detection in 20 years: A survey[J]. arXiv preprint arXiv:1905.05055, 2019.”与文献:“Wu X, Sahoo D, Hoi S C H. Recent advances in deep learning for object detection[J]. Neurocomputing, 2020.”);(2)Two-Stage目标检测算法,这类检测算法将检测问题划分为两个阶段,第一个阶段首先产生候选区域(Region Proposals),包含目标大概的位置信息,然后第二个阶段对候选区域进行分类和位置精修,这类算法的典型代表有R-CNN, Fast R-CNN, Faster R-CNN等。目标检测模型的主要性能指标是检测准确度和速度,其中准确度主要考虑物体的定位以及分类准确度。一般情况下,Two-Stage算法在准确度上有优势,而One-Stage算法在速度上有优势。

[0003] 近年来涌现出了一批旋转目标检测方法,主要是针对文字或遥感航拍图像,典型数据库为DOTA(参见文献:“Xia G S, Bai X, Ding J, et al. DOTA: A large-scale dataset for object detection in aerial images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3974-3983.”),典型算法有RRPN、RoI Transformer、BBAVectors、R3Det、CSL、SCRDet(参见文献:“Yang X, Yang J, Yan J, et al. Scrdet: Towards more robust detection for small, cluttered and rotated objects[C]//Proceedings of the IEEE International Conference on Computer Vision. 2019: 8232-8241.”)、S2A-Net(参见文献:“Han J, Ding J, Li J, et al. Align Deep Features for Oriented Object Detection[J]. arXiv preprint arXiv:2008.09397, 2020.”)等等,但现有算法都很难很好的解决旋转目标检测的几个难点,比如特征的对齐、旋转角度的精确回归,此外,现有算法还存在网络模型复杂、内存和计算资源需求大、建议锚框多且多无效导致的检测效率与精度低问题、多任务训练导致的模型污染目标定位精度低等问题。

发明内容

[0004] 为了解决现有技术中的上述问题,即为了解决航拍图像目标定位检测精度、识别精度及效率较低的问题,本发明第一方面,提出了一种插拔式航拍图像目标定位检测方法,该方法包括:

步骤S10,获取待定位检测的航拍图像,作为输入图像;

步骤S20,通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置;

步骤S30,基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

步骤S40,将各第二锚框映射至所述输入图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

[0005] 在一些优选的实施方式中,所述神经网络模型基于残差模块、第一卷积分支、第二卷积分支构建;所述第一卷积分支和第二卷积分支均包括一个 3×3 卷积层、一个 1×1 卷积层;其中,所述第一卷积分支用于生成目标中心点的热力图;所述第二个卷积分支用于生成目标中心点的偏移图。

[0006] 在一些优选的实施方式中,“通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置”,其方法为:

提取所述输入图像的特征,并分别通过第一卷积分支、第二卷积分支进行卷积处理,得到各目标的目标中心点的热力图、偏移图;

基于各目标的目标中心点的热力图,通过非极大值抑制方法NMS得到各目标的类别、中心位置;并根据所述目标中心点的偏移图对中心位置进行调整。

[0007] 在一些优选的实施方式中,所述神经网络模型其在训练过程中的损失函数为:

$$L = L_k + L_{off}$$

$$L_k = \frac{-1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}) & \text{if } Y_{xyc} = 1 \\ (1 - \hat{Y}_{xyc})^\beta (\hat{Y}_{xyc})^\alpha & \\ \log(1 - \hat{Y}_{xyc}) & \text{otherwise} \end{cases}$$

$$L_{off} = \frac{1}{N} \sum_p \left| \hat{o}_{\tilde{p}} - \left(\frac{p}{R} - \tilde{p} \right) \right|$$

其中, L 表示总损失, L_k 表示目标中心热力图损失, L_{off} 表示目标中心偏移损失, N 为目标中心点个数, α 、 β 为Focal Loss的超参, \hat{Y}_{xyc} 表示检测到的目标中心点置信度, Y_{xyc} 表示高斯核函数, p 表示每个目标类别的所有目标中心点, R 为输出步长, $\tilde{p} = \left\lfloor \frac{p}{R} \right\rfloor$, \hat{o} 为预测的中心点局部偏移。

[0008] 在一些优选的实施方式中,步骤S30中“对各第一锚框进行旋转处理”,其方法为:以各第一锚框的框中心为旋转中心,长边作为X轴,短边作为Y轴,在 $0-180^\circ$ 范围内每M度旋转一次;M为每一次旋转的角度,为正整数。

[0009] 本发明的第二方面,提出了一种插拔式航拍图像目标定位检测系统,该系统包括:图像获取模块、类别及位置获取模块、锚框预处理模块、边界框预测模块:

所述图像获取模块,配置为获取待定位检测的航拍图像,作为输入图像;

所述类别及位置获取模块,配置为通过预训练的神经网络模型获取所述输入图像

中的各目标的类别、中心位置；

所述锚框预处理模块，配置为基于各目标的类别、中心位置，结合预设第一关系，生成不同比例大小的锚框，作为第一锚框；并对各第一锚框进行旋转处理，得到对应的第二锚框；所述预设的第一关系为目标类别-设定长宽比例的映射关系；

所述边界框预测模块，配置为将各第二锚框映射至所述输入图像提取的特征图中，并通过感兴趣区域池化与回归处理，得到各目标对应的预测边界框。

[0010] 本发明的第三方面，提出了一种航拍图像目标类别识别系统，该系统包括：图像获取模块、类别及位置获取模块；

所述图像获取模块，配置为获取待识别的航拍图像，作为输入图像；

所述类别及位置获取模块，配置为通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置；

其中，所述神经网络模型基于残差模块、第一卷积分支、第二卷积分支构建；所述第一卷积分支和第二卷积分支均包括一个 3×3 卷积层、一个 1×1 卷积层；其中，所述第一卷积分支用于生成目标中心点的热力图；所述第二个卷积分支用于生成目标中心点的偏移图。

[0011] 本发明的第四方面，提出了一种航拍图像目标定位检测系统，该系统包括：提取模块、锚框预处理模块、边界框预测模块；

所述提取模块，配置为获取待目标定位检测的航拍图像，并提取该航拍图像中各目标的类别、中心位置；

所述锚框预处理模块，配置为基于各目标的类别、中心位置，结合预设第一关系，生成不同比例大小的锚框，作为第一锚框；并对各第一锚框进行旋转处理，得到对应的第二锚框；所述预设的第一关系为目标类别-设定长宽比例的映射关系；

所述边界框预测模块，配置为将各第二锚框映射至所述航拍图像提取的特征图中，并通过感兴趣区域池化与回归处理，得到各目标对应的预测边界框。

[0012] 本发明的第五方面，提出了一种插拔式航拍图像目标定位检测设备，包括：至少一个处理器；以及与至少一个所述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处理器执行的指令，所述指令用于被所述处理器执行以实现上述的插拔式航拍图像目标定位检测方法。

[0013] 本发明的第六方面，提出了一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有计算机指令，所述计算机指令用于被所述计算机执行以实现上述的插拔式航拍图像目标定位检测方法。

[0014] 本发明的有益效果：

本发明提高了航拍图像目标定位检测精度、识别精度及效率。

[0015] (1) 本发明神经网络模型只对目标中心位置及目标类别进行预测，没有其他回归预测分支，最大化的避免了模型污染以及特征不对齐造成的精度损失，也提高了预测效率，经实验验证，所涉及的模型较CenterNet等多任务回归模型（同时回归目标宽高等）可大大提升了目标中心点位置定位及类别识别的精度，较目前DOTA数据库竞赛中最好算法S2ANet的目标中心位置及类别预测精度都要高，且速度快。

[0016] (2) 本发明基于神经网络模型获取的目标中心点以及类别生成锚框，并结合

训练中目标的统计信息及目标类别信息,生成不同比例大小的锚框,不同于额外的RPN (Region Proposal Networks)网络提供的建议锚框,也不同于特征图每个像素遍历生成的密集锚框,不存在大量无效锚框问题,且锚框对感知的不同类别具有特定的统计设定,更接近真实框更容易实现精细回归。

[0017] (3)本发明中的深度神经网络模型、精细回归网络模块可针对不同的任务灵活“插拔”组合,充分有效地利用有限的无人机机载计算资源。

附图说明

[0018] 通过阅读参照以下附图所做的对非限制性实施例所做的详细描述,本申请的其他特征、目的和优点将会变得更明显。

[0019] 图1是本发明一种实施例的插拔式航拍图像目标定位检测方法的流程示意图;

图2是本发明一种实施例的插拔式航拍图像目标定位检测系统的结构示意图;

图3是本发明一种实施例的深度神经网络模型、精细回归网络模块的训练过程示意图;

图4是本发明一种实施例的航拍图像目标类别识别系统的结构示意图;

图5是本发明一种实施例的航拍图像目标定位检测系统的结构示意图;

图6是本发明一种实施例的通过深度神经网络模型获取的识别结果的示意图;

图7是本发明一种实施例的通过精细回归网络模块获取的检测结果的示意图。

具体实施方式

[0020] 为使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例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他实施例,都属于本发明保护的范围。

[0021] 下面结合附图和实施例对本申请作进一步的详细说明。可以理解的是,此处所描述的具体实施例仅用于解释相关发明,而非对该发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与有关发明相关的部分。

[0022] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0023] 本发明第一实施例的一种插拔式航拍图像目标定位检测方法,如图1所示,包括以下步骤:

步骤S10,获取待定位检测的航拍图像,作为输入图像;

步骤S20,通过预训练的深度神经网络模型获取所述输入图像中的各目标的类别、中心位置;

步骤S30,基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

步骤S40,将各第二锚框映射至所述输入图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

[0024] 为了更清晰地对本发明插拔式航拍图像目标定位检测方法进行说明,下面结合附

图对本发明方法一种实施例中各步骤进行展开详述。

[0025] 在下述实施例中,先对深度神经网络模型、精细回归网络模块的训练过程进行详述,再对通过插拔式航拍图像目标定位检测方法获取目标的预测边界框的过程进行详述。

[0026] 1、深度神经网络模型、精细回归网络模块的训练过程,如图3所示

步骤A10,深度神经网络模型训练

在本实施例中,利用航拍数据库训练深度神经网络模型。其中,深度神经网络模型主要包括BackBone部分(神经模型)和Neck部分,其中BackBone部分为残差模块,可以使用但不限于ResNet、DLA(Deep Layer Aggregation)、Hourglass等网络,Neck部分包含两个卷积分支,分别作为第一卷积分支、第二卷积分支,每个分支都是 3×3 卷积接 1×1 卷积,第一卷积分支生成目标中心点的热力图(Heatmap P)为 $\bar{W} \times \bar{H} \times C$ 大小的输出,其中 \bar{W} 与 \bar{H} 分别为最后一级特征图的宽和高,C为目标类别数目,通过NMS可得到目标的类别及中心点,第二卷积分支卷积生成目标中心点的偏移图(Offset O)为 $\bar{W} \times \bar{H} \times 2$ 大小的输出,对目标中心点进行精细调整。

[0027] 利用航拍数据库中的航拍图像以及标注的监督信息(目标中心点位置坐标及目标类别)对深度卷积神经网络模型进行训练,训练时的损失函数包括目标中心热图损失 L_k 及目标中心偏移损失 L_{off} ,即总的损失 L 为:

$$L = L_k + L_{off} \quad (1)$$

目标中心热图损失 L_k 使用基于Focal Loss的像素级逻辑回归损失,如公式(2)所示:

$$L_k = \frac{-1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}) & \text{if } Y_{xyc} = 1 \\ (1 - \hat{Y}_{xyc})^\beta (\hat{Y}_{xyc})^\alpha & \text{otherwise} \\ \log(1 - \hat{Y}_{xyc}) & \end{cases} \quad (2)$$

其中,预测 $\hat{Y}_{x,y,c} = 1$ 对应检测到的目标中心点,预测 $\hat{Y}_{x,y,c} = 0$ 对应背景,训练时对于标注的每个目标类别 c 的所有目标中心点 $p \in \mathcal{R}^2$,计算一个低分辨率的近似 $\tilde{p} = \lfloor \frac{p}{R} \rfloor$ (R 为输出步长),然后使用Gaussian(高斯)核函数 $Y_{xyc} = \exp(-\frac{(x-\tilde{p}_x)^2 + (y-\tilde{p}_y)^2}{2\sigma_p^2})$ 将标注的目标中心点生成到热力图 $Y \in [0,1]^{\frac{W}{R} \times \frac{H}{R} \times C}$ 中, N 为目标中心点个数, α 、 β 为Focal Loss的超参, \tilde{p}_x 、 \tilde{p}_y 表示 \tilde{p} 对应的坐标值。

[0028] 目标中心偏移损失主要用于解决在低分辨率上离散化造成的精度损失,如公式(3)所示:

$$L_{off} = \frac{1}{N} \sum_p \left| \hat{o}_{\tilde{p}} - \left(\frac{p}{R} - \tilde{p} \right) \right| \quad (3)$$

其中, $\hat{O} \in \mathcal{R}^{\frac{W}{R} \times \frac{H}{R} \times 2}$ 为预测的中心局部偏移, $\frac{W}{R}$ 、 $\frac{H}{R}$ 为 \bar{W} 、 \bar{H} 。

[0029] 训练后的深度神经网络模型可用于无人机目标快速定位打击等任务。

[0030] 步骤A20,精细回归网络模块训练

在本实施例中,精细回归网络模块(包括一个特征提取网络、感兴趣区域池化层ROI Pooling、全连接层FC256)可插于深度神经网络模型之后,实现快速的回归出目标的预测边界框。本发明利用目标中心点、目标类别、目标旋转框标注信息及训练图像,训练一个单独的精细回归网络模块。现有的算法生成锚框都是遍历整个特征图上的每个像素位置,对每个像素位置生成设定的几种比例、大小的锚框,生成的锚框极多而且对不同类别的目标生成锚框时并不区分,导致速度大大下降且有效锚框少,本发明训练时只在目标标注的中心点位置,利用训练集中目标的统计信息及目标标注的类别信息,对不同的目标类别生成特定比例不同大小的类别感知适应锚框。即基于深度神经网络模型获取的各目标的类别、中心位置,结合预设的目标类别-设定长宽比例的映射关系,生成不同比例大小的锚框(或简称为类别感知适应锚框),作为第一锚框。如图3中的大型交通工具(Large Vehicle)和小型交通工具(Small Vehicle)生成的锚框具有相应特定的长宽比例,不同于以往算法,不需要对每个像素位置生成多种不同长宽比例的锚框。

[0031] 对生成的类别感知适应锚框进行旋转,生成类别感知适应旋转锚框,作为第二锚框。具体为:以类别感知适应锚框中心为旋转中心长边作为X轴短边作为Y轴,0-180°范围内可每M度旋转一次生成不同方向的类别感知适应旋转锚框(因锚框具有类别感知特性,不在需要对每个点生成多种比例的锚框,只需对特定点特定类别生成特定比例锚框,锚框数大大降低。本发明中,M优选设置为10,在其他实施例中还可设定提取更多更精细的旋转框,比如每隔5°旋转生成36个不同方向的类别感知适应旋转锚框),提取第二锚框所对应的网络特征图(即将第二锚框映射至航拍图像的特征图中)并输入至ROI Pooling层进行池化(Pooling),图3中ROI Pooling优选采用3×3,实际使用时并不限于3×3,还可设计为5×5、7×7的ROI Pooling等,ROI Pooling后接全连接层FC-256进行回归,实现对最终目标旋转框(即边界框)的精细回归。回归后的边界框可支撑更多更高要求的任务。

[0032] 本发明中的深度神经网络模型、精细回归网络模块耦合度低,对不同的任务可灵活配置,即可以通过深度神经网络模型实现图像类别的识别,也可以通过单独通过训练好的精细回归网络模块实现边界框的预测。

[0033] 2、插拔式航拍图像目标定位检测方法

步骤S10,获取待定位检测的航拍图像,作为输入图像;

在本实施例中,获取待定位检测的航拍图像。

[0034] 步骤S20,通过预训练的深度神经网络模型获取所述输入图像中的各目标的类别、中心位置;

在本实施例中,通过上述训练好的深度神经网络模型提取所述输入图像的特征,并分别通过第一卷积分支、第二卷积分支进行卷积处理,得到各目标的目标中心点的热力图、偏移图。

[0035] 基于目标中心点的热力图,通过非极大值抑制方法NMS得到各目标的类别、中心位置,并根据目标中心点的偏移图对中心位置进行调整。对航拍图相中各目标的类别及中心位置的识别结果,如图6所示。

[0036] 步骤S30,提取所述输入图像的特征;基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应

的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

在本实施例中,基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框。以各第一锚框的框中心为旋转中心,长边作为X轴,短边作为Y轴,在0-180°范围内每M度旋转一次;M为每一次旋转的角度,为正整数。

[0037] 步骤S40,将各第二锚框映射至所述输入图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

[0038] 在本实施例中,提取输入图像的特征图,将各第二锚框映射至输入图像的特征图中,并通过ROI Pooling和FC-256层进行池化与回归处理,得到各目标对应的预测边界框。获取航拍图像各目标的预测边界框的预测结果,如图7所示。

[0039] 本发明第二实施例的一种插拔式航拍图像目标定位检测系统,如图2所示,该系统包括:图像获取模块101、类别及位置获取模块102、锚框预处理模块103、边界框预测模块104:

所述图像获取模块101,配置为获取待定位检测的航拍图像,作为输入图像;

所述类别及位置获取模块102,配置为通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置;

所述锚框预处理模块103,配置为基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

所述边界框预测模块104,配置为将各第二锚框映射至所述输入图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

[0040] 本发明第三实施例的一种航拍图像目标类别识别系统,如图4所示,该系统包括:图像获取模块201、类别及位置获取模块202;

所述图像获取模块201,配置为获取待识别的航拍图像,作为输入图像;

所述类别及位置获取模块202,配置为通过预训练的神经网络模型获取所述输入图像中的各目标的类别、中心位置;

其中,所述神经网络模型基于残差模块、第一卷积分支、第二卷积分支构建;所述第一卷积分支和第二卷积分支均包括一个 3×3 卷积层、一个 1×1 卷积层;其中,所述第一卷积分支用于生成目标中心点的热力图;所述第二个卷积分支用于生成目标中心点的偏移图。

[0041] 本发明第四实施例的一种航拍图像目标定位检测系统,如图5所示,该系统包括:提取模块301、锚框预处理模块302、边界框预测模块303;

所述提取模块301,配置为获取待目标定位检测的航拍图像,并提取该航拍图像中各目标的类别、中心位置;

所述锚框预处理模块302,配置为基于各目标的类别、中心位置,结合预设第一关系,生成不同比例大小的锚框,作为第一锚框;并对各第一锚框进行旋转处理,得到对应的第二锚框;所述预设的第一关系为目标类别-设定长宽比例的映射关系;

所述边界框预测模块303,配置为将各第二锚框映射至所述航拍图像提取的特征图中,并通过感兴趣区域池化与回归处理,得到各目标对应的预测边界框。

[0042] 所述技术领域的技术人员可以清楚的了解到,为描述的方便和简洁,上述描述的

系统具体的工作过程及有关说明,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0043] 需要说明的是,上述实施例提供的插拔式航拍图像目标定位检测系统/航拍图像目标类别识别系统/航拍图像目标定位检测系统,仅以上述各功能模块的划分进行举例说明,在实际应用中,可以根据需要而将上述功能分配由不同的功能模块来完成,即将本发明实施例中的模块或者步骤再分解或者组合,例如,上述实施例的模块可以合并为一个模块,也可以进一步拆分成多个子模块,以完成以上描述的全部或者部分功能。对于本发明实施例中涉及的模块、步骤的名称,仅仅是为了区分各个模块或者步骤,不视为对本发明的不当限定。

[0044] 本发明第五实施例的一种插拔式航拍图像目标定位检测设备,包括:至少一个处理器;以及与至少一个所述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处理器执行的指令,所述指令用于被所述处理器执行以实现上述的插拔式航拍图像目标定位检测方法。

[0045] 本发明的第六实施例的一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有计算机指令,所述计算机指令用于被所述计算机执行以实现上述的插拔式航拍图像目标定位检测方法。

[0046] 所述技术领域的技术人员可以清楚的了解,未描述的方便和简洁,上述描述的存储装置、处理装置的具体工作过程及有关说明,可以参考前述方法实例中的对应过程,在此不再赘述。

[0047] 需要说明的是,本申请上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存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储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,计算机可读存储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系统、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存储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系统、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。

[0048] 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,上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”语言或类似的程序设计语言。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)一连接到用户计算机,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)。

[0049] 附图中的流程图和框图,图示了按照本申请各种实施例的系统、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。也要注意的是,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。

[0050] 术语“第一”、“第二”等是用于区别类似的对象,而不是用于描述或表示特定的顺序或先后次序。

[0051] 术语“包括”或者任何其它类似用语旨在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备/装置不仅包括那些要素,而且还包括没有明确列出的其它要素,或者还包括这些过程、方法、物品或者设备/装置所固有的要素。

[0052] 至此,已经结合附图所示的优选实施方式描述了本发明的技术方案,但是,本领域技术人员容易理解的是,本发明的保护范围显然不局限于这些具体实施方式。在不偏离本发明的原理的前提下,本领域技术人员可以对相关技术特征作出等同的更改或替换,这些更改或替换之后的技术方案都将落入本发明的保护范围之内。

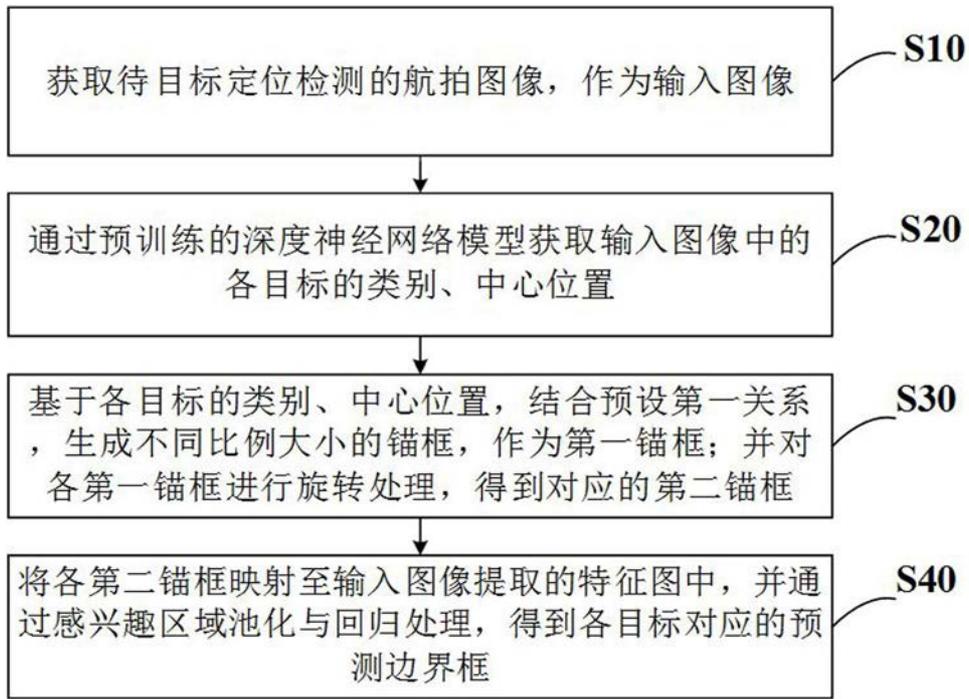


图 1



图 2

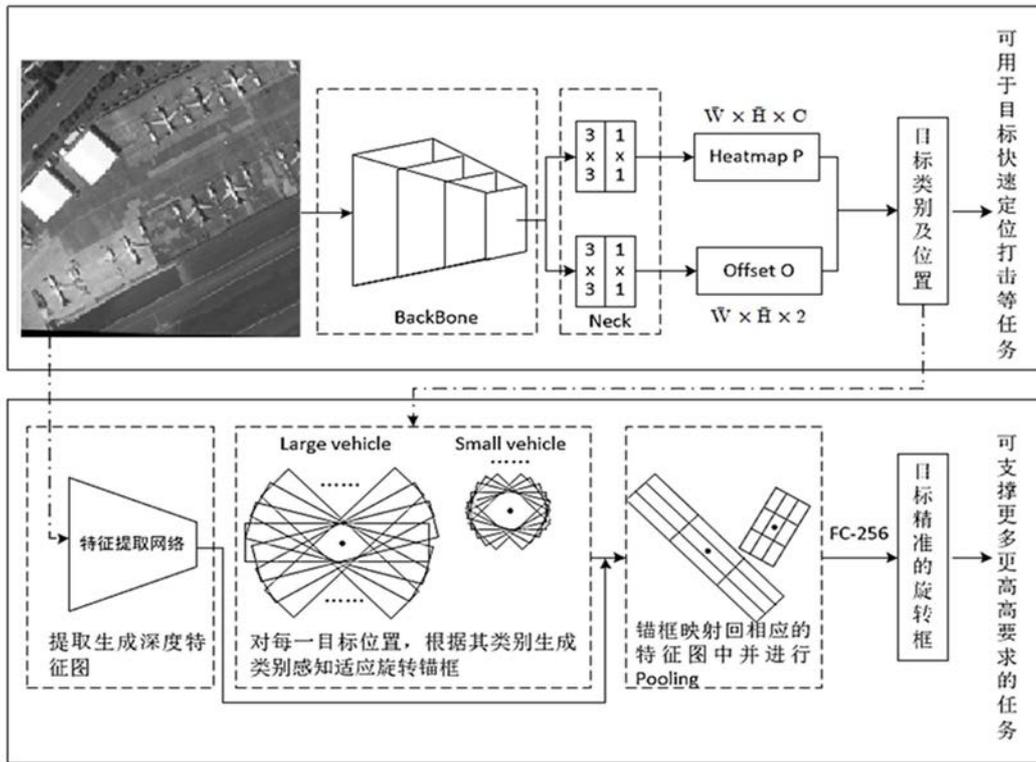


图 3



图 4



图 5



图 7