



(12) 发明专利申请

(10) 申请公布号 CN 114973014 A

(43) 申请公布日 2022.08.30

(21) 申请号 202210584494.7

G06N 3/04 (2006.01)

(22) 申请日 2022.05.27

G06N 3/08 (2006.01)

(71) 申请人 中国人民解放军战略支援部队信息工程大学

G06V 10/25 (2022.01)

G06V 10/40 (2022.01)

G06V 10/80 (2022.01)

地址 450000 河南省郑州市高新区科学大道62号

(72) 发明人 李润生 胡庆 潘超凡 牛朝阳 刘伟 许岩

(74) 专利代理机构 郑州大通专利商标代理有限公司 41111

专利代理师 周艳巧

(51) Int. Cl.

G06V 20/13 (2022.01)

G06V 10/764 (2022.01)

G06V 10/82 (2022.01)

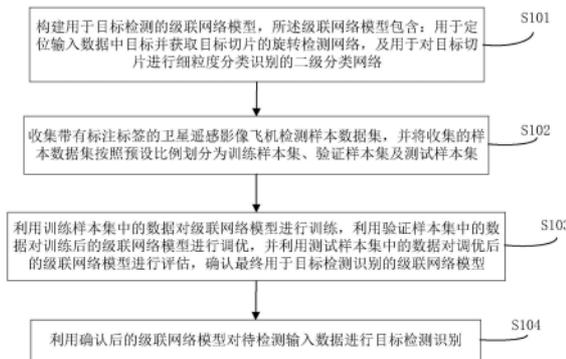
权利要求书2页 说明书12页 附图8页

(54) 发明名称

基于多网络级联的飞机目标细粒度检测方法及系统

(57) 摘要

本发明属于卫星遥感目标识别技术领域,特别涉及一种基于多网络级联的飞机目标细粒度检测方法及系统,通过构建用于目标检测的级联网络模型,所述级联网络模型包含:用于定位输入数据中目标并获取目标切片的旋转检测网络,及用于对目标切片进行细粒度分类识别的二级分类网络;收集带有标注标签的卫星遥感影像飞机检测样本数据集,并将收集的样本数据集按照预设比例划分为训练样本集、验证样本集及测试样本集,并分别对级联网络模型进行训练、调优和评估,确认最终用于目标检测识别的级联网络模型;利用级联网络模型对待检测输入数据进行目标检测识别。本发明将飞机检测和类型识别功能从单一网络中分离,便于充分利用公开数据集进行网络模型训练,分别提高检测和类型识别准确性。



1. 一种基于多网络级联的飞机目标细粒度检测方法,其特征在于,包含如下内容:

构建用于目标检测的级联网络模型,所述级联网络模型包含:用于定位输入数据中目标并获取目标切片的旋转检测网络,及用于对目标切片进行细粒度分类识别的二级分类网络;

收集带有标注标签的卫星遥感影像飞机检测样本数据集,并将收集的样本数据集按照预设比例划分为训练样本集、验证样本集及测试样本集;

利用训练样本集中的数据对级联网络模型进行训练,利用验证样本集中的数据对训练后的级联网络模型进行调优,并利用测试样本集中的数据对调优后的级联网络模型进行评估,确认最终用于目标检测识别的级联网络模型;

利用确认后的级联网络模型对待检测输入数据进行目标检测识别。

2. 根据权利要求1所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,旋转检测网络中,利用YOLOv5作为框定输入数据目标位置并获取检测框中目标切片的主干网络,并在主干网络输出层上添加用于角度分类回归的角度预测回归器。

3. 根据权利要求2所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,主干网络中采用六参数定义水平回归框,具体定义内容描述如下:(category, x, y, w, h, θ),其中,category表示类别序号,(x, y)表示检测框中心点坐标,w表示矩形检测框较长边,h表示矩形检测框短边, θ 表示x轴沿顺时针方向旋转遇到矩形检测框最长边所经过的角度。

4. 根据权利要求3所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,旋转检测网络中,使用高斯窗口函数进行标签平滑处理,以使网络衡量预测标签和真实标签之间的角度距离;并利用纵横比感知权重函数来约束检测框在旋转角度下的纵横比。

5. 根据权利要求4所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,高斯窗口函数表示为: $Guass(x) = \begin{cases} e^{-(x-\theta)^2/2\sigma^2}, & \theta-r < x < \theta+r \\ 0, & \text{其他} \end{cases}$, 其中, σ 为窗函数方差,r为窗函数半径;纵横比感知权重函数表示为:

$$w(\Delta\theta) = \begin{cases} 1 - \cos^{2\alpha}(2\Delta\theta) = 1 - \cos^{2\alpha}(2(\theta_{gt} - \theta_{pred})) \cdot \frac{w_{gt}}{h_{gt}}, & (h_{gt}/w_{gt}) < r \\ 1, & \text{其他} \end{cases}$$
, 其中, h_{gt} 和 w_{gt} 分别为

检测框长边和短边,r为纵横比阈值, α 为调节权重窗口超参数, $\Delta\theta$ 表示预测角度值与真实值之差, θ_{gt} 表示目标角度真实值, θ_{pred} 表示角度预测值。

6. 根据权利要求1所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,旋转检测网络的目标损失函数表示为: $Loss = \lambda_1 L_{box} + \lambda_2 L_{cls} + \lambda_3 L_{conf} + \lambda_4 L_0$,其中, $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 分别为位置损失 L_{box} ,分类损失 L_{cls} ,置信度损失 L_{conf} 和角度损失 L_0 的超参数。

7. 根据权利要求1所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,二级分类网络采用EfficientNet-v2s网络结构,并该网络结构的池化层输出侧增加由多个子分类器构建集成学习模块和对提取特征进行约束的Fisher正则化判别层。

8. 根据权利要求7所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,二级分类网络中,利用EfficientNet-v2s网络结构提取一级特征,并按照设定比率随机选取所有一级特征中的特征数据,将选取的特征数据进行融合产生二级特征;利用子分类器对

每个二级特征进行预分类,并将所有子分类器的预分类结果通过平均加权生成最终分类结果。

9.根据权利要求7或8所述的基于多网络级联的飞机目标细粒度检测方法,其特征在于,二级分类网络的目标损失函数表示为: $J = \min \left(\lambda_0 J_0(X, Y) + \lambda_1 J_1(X) + \frac{\lambda_2}{2} \|W\|_F^2 \right)$,其中, $\lambda_0, \lambda_1, \lambda_2$ 分别为分类损失函数项 $J_0(X, Y)$,提取特征施加Fisher判别正则化函数项 $J_1(X)$ 及权重衰减函数项 $\|W\|_F^2$ 的超参数, X, Y 分别为输入数据及其对应标签。

10.一种基于多网络级联的飞机目标细粒度检测系统,其特征在于,包含:模型构建模块、样本收集模块、模型训练模块及目标识别模块,其中,

模型构建模块,用于构建用于目标检测的级联网络模型,所述级联网络模型包含:用于定位输入数据中目标并获取目标切片的旋转检测网络,及用于对目标切片进行细粒度分类识别的二级分类网络;

样本收集模块,用于收集带有标注标签的卫星遥感影像飞机检测样本数据集,并将收集的样本数据集按照预设比例划分为训练样本集、验证样本集及测试样本集;

模型训练模块,用于利用训练样本集中的数据对级联网络模型进行训练,利用验证样本集中的数据对训练后的级联网络模型进行调优,并利用测试样本集中的数据对调优后的级联网络模型进行评估,确认最终用于目标检测识别的级联网络模型;

目标识别模块,用于利用确认后的级联网络模型对待检测输入数据进行目标检测识别。

基于多网络级联的飞机目标细粒度检测方法及系统

技术领域

[0001] 本发明属于卫星遥感目标识别技术领域,特别涉及一种基于多网络级联的飞机目标细粒度检测方法及系统。

背景技术

[0002] 飞机的检测和识别是光学遥感图像解释领域的一项重要任务。随着遥感影像分辨率的提高,可以捕捉到目标越来越多的色彩、纹理和空间信息,使得进一步准确分析判读飞机的属性和型号成为现实。遥感影像飞机的细粒度检测包含飞机目标的方位预测和型号细粒度识别,由于成像背景、阴影、光照、分辨率、姿态、视点、飞机涂装和不同类型飞机外型相似等复合因素,飞机的型号识别一直是一项充满挑战性的任务。

[0003] 遥感影像飞机目标的细粒度检测存在以下几个难点:一是飞机目标的类内多样性和类间相似性给型号识别带来挑战,比如以波音707为原型改造的运输机C-135、电子侦察机RC-135和空中加油机KC-135在外形上就极具相似性,需要通过识别特定部件的细小差别才能区分,对网络的鉴别能力提出了更高的要求,而同一类型飞机又会由于颜色、成像角度、背景等因素导致差异性。二是数据集稀缺,现有开源的包含飞机型号的目标检测数据集很少,多数机型的数量仅有几十架,类别不平衡、有限的样本和低质量影像给训练鲁棒的目标检测网络带来挑战,单一网络泛化性能不足,在同时实现目标定位和机型识别时的精度有待提升。

发明内容

[0004] 为此,本发明提供一种基于多网络级联的飞机目标细粒度检测方法及系统,将飞机检测和类型识别的功能从单一网络中分离出来,可以分别提高检测和类型识别的准确性,便于充分利用公开数据集进行网络模型训练,可适用于不同场景和分辨率下目标型号分类识别。

[0005] 按照本发明所提供的设计方案,提供一种基于多网络级联的飞机目标细粒度检测方法,包含如下内容:

[0006] 构建用于目标检测的级联网络模型,所述级联网络模型包含:用于定位输入数据中目标并获取目标切片的旋转检测网络,及用于对目标切片进行细粒度分类识别的二级分类网络;

[0007] 收集带有标注标签的卫星遥感影像飞机检测样本数据集,并将收集的样本数据集按照预设比例划分为训练样本集、验证样本集及测试样本集;

[0008] 利用训练样本集中的数据对级联网络模型进行训练,利用验证样本集中的数据对训练后的级联网络模型进行调优,并利用测试样本集中的数据对调优后的级联网络模型进行评估,确认最终用于目标检测识别的级联网络模型;

[0009] 利用确认后的级联网络模型对待检测输入数据进行目标检测识别。

[0010] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,旋转检测网

络中,利用YOLOv5作为框定输入数据目标位置并获取检测框中目标切片的主干网络,并在主干网络输出层上添加用于角度分类回归的角度预测回归器。

[0011] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,主干网络中采用六参数定义水平回归框,具体定义内容描述如下: $(category, x, y, w, h, \theta)$, 其中, $category$ 表示类别序号, (x, y) 表示检测框中心点坐标, w 表示矩形检测框较长边, h 表示矩形检测框短边, θ 表示 x 轴沿顺时针方向旋转遇到矩形检测框最长边所经过的角度。

[0012] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,旋转检测网络中,使用高斯窗口函数进行标签平滑处理,以使网络衡量预测标签和真实标签之间的角度距离;并利用纵横比感知权重函数来约束检测框在旋转角度下的纵横比。

[0013] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,高斯窗口函数表示为: $Guass(x) = \begin{cases} e^{-(x-\theta)^2/2\sigma^2}, & \theta-r < x < \theta+r \\ 0, & \text{其他} \end{cases}$, 其中, σ 为窗函数方差, r 为窗函数半径;

纵横比感知权重函数表示为:

[0014] $w(\Delta\theta) = \begin{cases} 1 - \cos^{2\alpha}(2\Delta\theta) = 1 - \cos^{2\alpha}(2(\theta_{gt} - \theta_{pred})) \cdot \frac{w_{gt}}{h_{gt}}, & (h_{gt}/w_{gt}) < r \\ 1, & \text{其他} \end{cases}$, 其中, h_{gt} 和

w_{gt} 分别为检测框长边和短边, r 为纵横比阈值, α 为调节权重窗口超参数, $\Delta\theta$ 表示预测角度值与真实值之差, θ_{gt} 表示目标角度真实值, θ_{pred} 表示角度预测值。

[0015] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,旋转检测网络的目标损失函数表示为: $Loss = \lambda_1 L_{box} + \lambda_2 L_{cls} + \lambda_3 L_{conf} + \lambda_4 L_\theta$, 其中, $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 分别为位置损失 L_{box} , 分类损失 L_{cls} , 置信度损失 L_{conf} 和角度损失 L_θ 的超参数。

[0016] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,二级分类网络采用EfficientNet-v2s网络结构,并该网络结构的池化层输出侧增加由多个子分类器构建集成学习模块和对提取特征进行约束的Fisher正则化判别层。

[0017] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,二级分类网络中,利用EfficientNet-v2s网络结构提取一级特征,并按照设定比率随机选取所有一级特征中的特征数据,将选取的特征数据进行融合产生二级特征;利用子分类器对每个二级特征进行预分类,并将所有子分类器的预分类结果通过平均加权生成最终分类结果。

[0018] 作为本发明基于多网络级联的飞机目标细粒度检测方法,进一步地,二级分类网络的目标损失函数表示为: $J = \min \left(\lambda_0 J_0(X, Y) + \lambda_1 J_1(X) + \frac{\lambda_2}{2} \|W\|_F^2 \right)$, 其中, $\lambda_0, \lambda_1, \lambda_2$ 分别为

分类损失函数 $J_0(X, Y)$, 提取特征施加Fisher判别正则化函数 $J_1(X)$ 及权重衰减项函数 $\|W\|_F^2$ 的超参数, X, Y 分别为输入数据及其对应标签。

[0019] 进一步地,本发明还提供一种基于多网络级联的飞机目标细粒度检测系统,包含:模型构建模块、样本收集模块、模型训练模块及目标识别模块,其中,

[0020] 模型构建模块,用于构建用于目标检测的级联网络模型,所述级联网络模型包含:用于定位输入数据中目标并获取目标切片的旋转检测网络,及用于对目标切片进行细粒度

分类识别的二级分类网络；

[0021] 样本收集模块,用于收集带有标注标签的卫星遥感影像飞机检测样本数据集,并将收集的样本数据集按照预设比例划分为训练样本集、验证样本集及测试样本集；

[0022] 模型训练模块,用于利用训练样本集中的数据对级联网络模型进行训练,利用验证样本集中的数据对训练后的级联网络模型进行调优,并利用测试样本集中的数据对调优后的级联网络模型进行评估,确认最终用于目标检测识别的级联网络模型；

[0023] 目标识别模块,用于利用确认后的级联网络模型对待检测输入数据进行目标检测识别。

[0024] 本发明的有益效果：

[0025] 本发明将飞机检测和类型识别的功能从单一网络中分离出来,可以分别提高检测和类型识别的准确性；另,因包含飞机型号的公开数据集较少,制作数据集耗时费力,该级联网络可以充分利用公开数据集训练鲁棒的飞机检测网络,增强在航空和卫星影像中的通用性,保证飞机检测的召回率；然后只需要采集一些各类飞机切片,对其进行缩放、雾化、旋转、亮度变化、对比度变化、场景迁移等数据增强操作,就可以生成大量的不同分辨率、不同姿势、不同光线、不同背景下的飞机切片,保证二级分类网络可以适应不同场景和分辨率下的机型分类,如果后续需要增加新的目标类别,只需要对二级分类网络进行微调或增量学习就能够很方便应用；解决光学遥感影像中的多方向密集排列、多种型号飞机目标的方向检测和型号识别困难问题。并进一步结合数据增强技术,通过集成学习和Fisher判别正则化改进分类网络,提升机型分类精度,使得本案方案在整体检测精度方面具有更大优势,具有较好的应用前景。

附图说明：

[0026] 图1为实施例中飞机目标细粒度检测方法流程示意；

[0027] 图2为实施例中级联网络模型结构示意；

[0028] 图3为实施例中旋转检测网络结构示意；

[0029] 图4为实施例中PoA和EoE问题示意；

[0030] 图5为实施例中长边定义法示意；

[0031] 图6为实施例中类正方形目标IoU精度和损失值不一致示意；

[0032] 图7为实施例中二级分类网络结构示意；

[0033] 图8为实施例中样本标签可视化示意；

[0034] 图9为实施例中选取部分目标样例示意；

[0035] 图10为实施例中数据增广方法示意；

[0036] 图11为实施例中DOTA测试集上的检测结果示意；

[0037] 图12为实施例中改进网络在测试集上的混淆矩阵示意；

[0038] 图13为实施例中单级网络检测结果混淆矩阵示意；

[0039] 图14为实施例中对比算法对各类目标检测精度的柱状图示意；

[0040] 图15为实施例中对比算法在不同场景下的检测结果示意；

[0041] 图16为实施例中本案算法在未知机场的检测结果显示。

具体实施方式：

[0042] 为使本发明的目的、技术方案和优点更加清楚、明白，下面结合附图和技术方案对本发明作进一步详细的说明。

[0043] 本发明实施例，参见图1所示，提供一种基于多网络级联的飞机目标细粒度检测方法，包含如下内容：

[0044] S101、构建用于目标检测的级联网络模型，所述级联网络模型包含：用于定位输入数据中目标并获取目标切片的旋转检测网络，及用于对目标切片进行细粒度分类识别的二级分类网络；

[0045] S102、收集带有标注标签的卫星遥感影像飞机检测样本数据集，并将收集的样本数据集按照预设比例划分为训练样本集、验证样本集及测试样本集；

[0046] S103、利用训练样本集中的数据对级联网络模型进行训练，利用验证样本集中的数据对训练后的级联网络模型进行调优，并利用测试样本集中的数据对调优后的级联网络模型进行评估，确认最终用于目标检测识别的级联网络模型；

[0047] S104、利用确认后的级联网络模型对待检测输入数据进行目标检测识别。

[0048] 本案实施例中，将飞机检测和类型识别的功能从单一网络中分离出来，可以分别提高检测和类型识别的准确性，以提升目标检测识别效果。

[0049] 为提高遥感影像飞机目标细粒度检测的水平，可利用构建的大型光学卫星遥感影像飞机检测数据集FDDA来作为样本数据。FDDA数据能够获取自高分、WorldView、Quickbird等国内外遥感卫星，拍摄地区包括各个区域的机场，影像空间分辨率介于1m~2m之间，可通过采集不同季节、不同视点、不同光照条件下的影像数据，以保证数据集的多样性。

[0050] FDDA数据集包含2 244幅尺寸为800*800像素~13 680*8 392像素不等的卫星遥感影像，可按照6:2:2的比例划分为训练集、验证集和测试集，分别包含1 346、449、449张影像，存储格式为.tif，数据量为46.2GB。经过专家判读，数据集中飞机目标实例总数为17 506个，包括轰炸机、预警机、运输机、加油机、战斗机、客机和直升机7个大类，根据具体型号又可分为13个小类，表1中展示了数据集的具体分类情况以及各类别数量统计，各类目标数量分布稍微有些不均衡，最大相差约8倍。部分类的目标个体实例，除了一些由民航改装的加油机和侦察机需要通过飞机翼型、机长翼展比等非常细微的特征点去辨别外，其他型号在外形布局上具有比较明显的区分特征。

[0051] 表1 FDDA数据集包含的飞机类别

大类	型号	数量
轰炸机	B-1B	1218
	B-52	2492
预警机	E-3	1220
	E-8	452
加油机	KC-10	1414
	KC-135	3655
运输机	C-130	1470
	C-17	985
战斗机	F-15	2042
	F-16	586
	F-22	556
客机	波音-737	1216
直升机	直升机	708

[0052] 所有目标实例可通过`rolabelImg`进行标注,标注样式可选择为旋转框,标注信息为 $(cx, cy, w, h, angle)$,分别表示中心点坐标 (cx, cy) 、标注框的宽和高 (w, h) 和标注框顺时针旋转的角度,角度范围为 $(0, \pi)$ 。

[0054] 本案实施例中,参见图2所示,为了实现对影像中的飞机细粒度检测,级联网络首先使用旋转检测算法从影像中框定飞机的位置,然后,切割检测框中的飞机目标,通过以EfficientNetV2-S为主体的细粒度分类网络实现快速精准的分类。

[0055] 进一步地,本案实施例的旋转检测网络中,利用YOLOv5作为框定输入数据目标位置并获取检测框中目标切片的主干网络,并在主干网络输出层上添加用于角度分类回归的角度预测回归器。

[0056] 基于旋转框的飞机检测算法可以有效地减少目标框中的冗余背景信息,实现更准确的定位预测。如图3所示的旋转检测网络,包括预测参数、损失函数等,以确保模型的功能实现和正确的收敛,主干网络采用YOLOv5,在检测输出阶段,添加角度预测模块,并修改相应的损失函数,利用纵横比感知权重函数来解决类正方形目标带来的角度损失突变问题,以实现更好的旋转检测功能。

[0057] 进一步地,本案实施例的旋转检测网络中,使用高斯窗口函数进行标签平滑处理,以使网络衡量预测标签和真实标签之间的角度距离;并利用纵横比感知权重函数来约束检测框在旋转角度下的纵横比。

[0058] 定向检测框参数化的主流方法主要有八参数定义法 $(x_a, y_a, x_b, y_b, x_c, y_c, x_d, y_d)$ 和五参数定义法 (x, y, w, h, θ) ,鉴于YOLOv5使用中心坐标和宽高 (x, y, w, h) 定义水平回归框。本案实施例中,增加一个角度参数 θ 来实现定向水平框的设定。但是五参数定义法存在损失不连续的问题,主要来源于角度的周期性(periodicity of angular, PoA)和边缘的可交换性(exchangeability of edges, EoE),如图4所示,浅灰是标签位置,角度为 0° ,当浅灰框逆时针旋转 2° 时到达黑色框位置,此时检测框的角度为 2° ;当浅灰框顺时针旋转 2° 时到达深灰框的位置,此时检测框的角度为 178° ,这两种旋转方式的实际角度误差值很小,损失值却产生了一个周期的突变,增加了稳定收敛的难度;(b)表示在预测目标框时,发生了宽高互换现象,OpenCV定义法规定的角度值是指目标框与x轴所成的锐角,范围为 $[-90^\circ, 0^\circ)$,由于理想的回归方式超出了角度设定范围,需要顺时针旋转至过渡框,再进行边的交换得到最

终预测框,这样的回归方式导致损失值较大。

[0059] 为了解决以上两个问题,可把角度回归问题转化为分类问题,通过限制预测结果的范围来消除边界损失突变问题,并使用长边定义法进行角度标记,如图5所示,数据标签可采用(category, x, y, w, h, θ)六参数形式,其中category表示类别序号, (x, y)表示检测框的中心点坐标, w表示检测框较长的那一边, h表示另一条短边, θ 表示x轴沿顺时针方向旋转到检测框最长边所经过的角度,范围是 $[0^\circ, 179^\circ]$ 。进一步,可将角度回归分为180类,间隔 1° 划分为一类,这样划分虽然无法预测出 0.5° 中间值,但对最终精度影响很小。

[0060] 由于角度预测转换成分类回归之后,角度真值就变成了离散的独热码,所以在角度分类损失计算时,传统的交叉熵损失函数并不能真实反映角度间的距离,比如当真实值为 0° 时,预测成 1° 和 90° 时候的损失值可能是一样的,但现实中应该允许预测值接近真实值,即对于真实角度为 0° 时,预测成 179° 时或者 1° 是能够接受的。所以在交叉熵计算损失之前,可使用高斯窗口函数进行标签平滑处理,使得模型可以衡量预测标签和真实标签之间的角度距离,即在一定范围内,越靠近真实值的预测值的损失值越小,以此实现更加鲁棒的角度预测。高斯窗口函数表达如下所示:

$$[0061] \quad \text{Guass}(x) = \begin{cases} e^{-(x-\theta)^2/2\sigma^2}, & \theta-r < x < \theta+r \\ 0, & \text{其他} \end{cases}$$

[0062] 式中: σ 是窗函数的方差; r 是窗函数的半径, σ 默认为6; θ 表示目标框的真实角度。

[0063] “角度分类+长边定义法”较好地解决了边界损失不连续问题,但是无法较好地定义一个类正方形目标,如图6所示,类正方形目标两边长近似相等,长边和短边的预测顺序容易相反,而且位置损失误差依然很小,(c)表示两个相同的目标框在不同旋转角度下的交并比曲线,每一条曲线代表一类纵横比,可以看出,当目标框旋转 90° 时,目标的长宽比越小,交并比越接近1,但是这两个目标框的预测角度却相差约 90° ,也就是目标IoU的评估和角度损失之间存在不一致性。为此,本案实施例中,利用纵横比感知权重函数(Angle Ratio Sensitive Weighting Function, ARSWF),具体公式如下:

$$[0064] \quad w(\Delta\theta) = \begin{cases} 1 - \cos^{2\alpha}(2\Delta\theta) = 1 - \cos^{2\alpha}(2(\theta_{gt} - \theta_{pred})) \cdot \frac{w_{gt}}{h_{gt}}, & (h_{gt}/w_{gt}) < r \\ 1, & \text{其他} \end{cases}$$

[0065] 式中: h_{gt} 和 w_{gt} 是真实框的长边和短边; r 是纵横比阈值,可从图6的(c)中选取 $r=1.1$; α 是调节权重窗口的超参数,默认取20。该函数可以灵活地针对不同纵横比的目标调整训练策略,以 $\frac{\pi}{2}$ 为周期抑制类正方形目标的角度损失,缓解不一致性,促进网络收敛。

[0066] 旋转检测网络将损失函数分为位置损失(L_{box})、分类损失(L_{cls})、置信度损失(L_{conf})和角度损失(L_θ),定义如下:

$$[0067] \quad \text{Loss} = \lambda_1 L_{box} + \lambda_2 L_{cls} + \lambda_3 L_{conf} + \lambda_4 L_\theta$$

[0068] 式中,超参数 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 控制这四项损失的平衡,默认设置为 $\{1, 0.5, 1, 1\}$ 。

[0069] 设定 (x, y, w, h, θ) 、 $(\hat{x}, \hat{y}, \hat{w}, \hat{h}, \hat{\theta})$ 分别代表真实标注框和预测框的参数,位置回归损失 L_{box} 考虑覆盖面积、中心点距离和长宽比这三个因素,使用CIoU计算,如下所示:

$$[0070] \quad L_{box} = 1 - IOU + \frac{d^2([x, y], [\hat{x}, \hat{y}])}{c^2} + \frac{v^2}{1 - IOU + v}.$$

[0071] 式中, IOU表示真实标注框和预测框的交并比; $d([x, y], [\hat{x}, \hat{y}])$ 表示两中心点坐标的欧氏距离; c 表示标注框和预测框最小外接矩形的对角线长度;

$v = \frac{4}{\pi^2} \left(\arctan \frac{w}{h} - \arctan \frac{\hat{w}}{\hat{h}} \right)^2$ 表示标注框与预测框长宽比的距离。

[0072] 可使用多分类交叉熵计算分类损失 L_{cls} 和角度损失 L_θ ,如下所示:

$$[0073] \quad L_{cls} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C \left[I\{\hat{c}_{ij} = True\} \cdot \ln(p_{ij}) + I\{\hat{c}_{ij} \neq True\} \cdot \ln(1 - p_{ij}) \right];$$

$$[0074] \quad L_\theta = -\frac{1}{N} \sum_{i=1}^N \left[\left(\sum_{j=1}^C \left[I\{\hat{c}_{ij} = True\} \cdot \ln(p_{ij}) + I\{\hat{c}_{ij} \neq True\} \cdot \ln(1 - p_{ij}) \right] \right) \cdot w(\Delta\theta) \right].$$

[0075] 式中, N 代表检测框的总数; C 代表类别数,在角度损失中即为180类; I 是符号函数,当第 i 个检测框的真实类别为 j 类时,值为1,否则为0; p_{ij} 表示网络输出的类别预测值; $w(\Delta\theta)$ 代表纵横比感知权重函数。

[0076] 置信度损失函数设置如下所示:

$$[0077] \quad L_{conf} = -\frac{1}{N} \sum_{i=1}^N \left[I\{\hat{c}_i = True\} \cdot CIoU \cdot \ln(conf_i) + I\{\hat{c}_i \neq True\} \cdot \ln(1 - conf_i) \right].$$

[0078] 式中, N 代表检测框的总数; $conf_i$ 表示置信度的预测值,对检测框为正样本的置信度真值不直接赋1,而是计算对应检测框与真实框的CIoU,使用CIoU作为该检测框的置信度标签,对检测框为负样本的置信度真值还是直接赋0,这样一来,标签值的大小与检测框、真实框的重合度有关,重合程度越高则标签值越大;若检测框为正样本,则 $I\{\hat{c}_i = True\} = 1$,否则 $I\{\hat{c}_i \neq True\} = 1$ 。

[0079] 进一步地,本案实施例中,二级分类网络采用EfficientNet-v2s网络结构,并该网络结构的池化层输出侧增加由多个子分类器构建集成学习模块和对提取特征进行约束的Fisher正则化判别层。二级分类网络中,利用EfficientNet-v2s网络结构提取一级特征,并按照设定比率随机选取所有一级特征中的特征数据,将选取的特征数据进行融合产生二级特征;利用子分类器对每个二级特征进行预分类,并将所有子分类器的预分类结果通过平均加权生成最终分类结果。

[0080] 光学遥感影像中目标的类内相异性、类间相似性以及尺寸方向变换等情况降低了目标检测器的性能,不同型号飞机目标之间尺寸接近、形状类似,类间特征差别较小,难以精细识别;同型号飞机之间受可变掠翼张角不同、涂装、自身阴影等影响,容易产生错误分类。本案实施例中,基于EfficientNetV2-S模型利用一个轻量的细粒度分类子网络,通过利用多个子分类器构建集成学习模块、引入Fisher判别正则化对CNN提取的特征施加约束来提升分类效果,整体网络架构如图7所示,分类网络的主体结构是EfficientNetV2-S,在池化层后,增加集成学习模块和Fisher正则化函数。

[0081] EfficientNetV2是谷歌大脑提出的一种新型高效分类网络。它基于感知神经架构搜索技术,获得一组最优的复合系数,将MBConv、fusion-MBConv、网络深度和宽度等进行缩

放和组合。通过限制输入图像的最大尺寸和使用渐进学习方案,实现了参数效率和分类精度之间的完美平衡。考虑到网络复杂度对检测时间的影响,本案实施例中,采用参数最少的EfficientnetV2-S网络,在池化层后,构建多个子分类器进行集成学习,并增加Fisher正则化约束以提高分类效果。

[0082] 在机器学习中,通过构建多组基分类器组成一个预测模型的元算法,以此来增强模型的稳定性和鲁棒性。本案实施例中,增加的集成学习模块使用多个子分类器实现优势互补,试图通过群体决策来提升模型学习能力,解决单一分类器性能不足的问题,各子分类器之间差异性越大,集成效果越好。基于EfficientNetV2-S网络提取一级特征,可按照设定比率60%对所有的一级特征进行随机选取,融合产生新的二级特征,给每个二级特征分别训练一个子分类器进行预分类,最后平均加权所有子分类器的预分类结果,生成最终的分类结果,如下式所示:

$$[0083] \quad H(x) = \frac{1}{T} \sum_{i=1}^T h_i(x).$$

[0084] 式中:T表示子分类器总个数,默认T=7; $h_i(x)$ 表示第i个子分类器的输出结果。

[0085] 其次,通过优化一个新的目标函数对CNN提取的一级特征施加Fisher判别准则来训练Fisher判别层,使目标分类具有较小的类内离散和较大的类间离散度,以提升细粒度分类的精度。给定一组训练样本 $X = \{x_1, x_2, x_3, \dots\}$ 及其相应标签 $Y = \{y_{x_i}, x_i \in X\}$,新的目标函数表达如下:

$$[0086] \quad J = \min \left(\lambda_0 J_0(X, Y) + \lambda_1 J_1(X) + \frac{\lambda_2}{2} \|W\|_F^2 \right).$$

[0087] 式中: $\lambda_0, \lambda_1, \lambda_2$ 是衡量各项重要性的超参数,这里分别设置为[1, 0.005, 0.0005]。第一项是分类损失函数,它是使用标签平滑-负对数似然准则定义的。其目的是最小化给定训练样本的分类误差,计算如下:

$$[0088] \quad J_0(X, Y) = - \sum_{i=1}^K p_i \ln \frac{\exp(O_s(x_i))}{\sum_{j=1}^K \exp(O_s(x_j))}$$

[0089] 式中:K表示总类别数, $O_s(x_i)$ 表示第i类目标的softmax输出,

$$p_i = \begin{cases} 1 - \varepsilon & , \text{if } y = y_i \\ \varepsilon / (K - 1) & , \text{others} \end{cases} \text{表示平滑处理后的标签值, } \varepsilon \text{ 是一个很小的常数, 设为 } 0.1.$$

[0090] 第二项 $J_1(X)$ 是对训练样本提取的特征施加Fisher判别正则化,目的是最小化类内离散度($S_W(X)$),最大化类间离散度($S_B(X)$),对于训练样本 $x \in X$, $S_W(X)$ 和 $S_B(X)$ 定义如下:

$$[0091] \quad S_W(X) = \sum_{k=1}^K \sum_{x_i \in X} (O(x_i) - m_k)(O(x_i) - m_k)^T;$$

$$[0092] \quad S_B(X) = \sum_{k=1}^K n_k (m_k - m)(m_k - m)^T.$$

[0093] 式中: $O(x)$ 表示池化层后输出的一级特征, n_k 表示第k类目标的样本总数, m_k 表示第k类目标的平均特征值, m 表示全部训练样本的平均特征值。判别正则化 $J_1(X)$ 定义如下:

[0094] $J_1(X) = \text{tr}(S_w(X)) - \text{tr}(S_B(X))$.

[0095] 式中： $\text{tr}(X)$ 表示矩阵X的对角线总和。第三项是权重衰减项，降低网络过拟合的可能性。

[0096] 进一步地，基于上述的方法，本发明实施例还提供一种基于多网络级联的飞机目标细粒度检测系统，包含：模型构建模块、样本收集模块、模型训练模块及目标识别模块，其中，

[0097] 模型构建模块，用于构建用于目标检测的级联网络模型，所述级联网络模型包含：用于定位输入数据中目标并获取目标切片的旋转检测网络，及用于对目标切片进行细粒度分类识别的二级分类网络；

[0098] 样本收集模块，用于收集带有标注标签的卫星遥感影像飞机检测样本数据集，并将收集的样本数据集按照预设比例划分为训练样本集、验证样本集及测试样本集；

[0099] 模型训练模块，用于利用训练样本集中的数据对级联网络模型进行训练，利用验证样本集中的数据对训练后的级联网络模型进行调优，并利用测试样本集中的数据对调优后的级联网络模型进行评估，确认最终用于目标检测识别的级联网络模型；

[0100] 目标识别模块，用于利用确认后的级联网络模型对待检测输入数据进行目标检测识别。

[0101] 为验证本案方案有效性，下面结合试验数据做进一步解释说明：

[0102] 实验采用自建数据集和“天智杯”平台的飞机智能检测识别数据。其中，“天智杯”数据集共有430幅影像，划分训练集308幅，测试集122幅，尺寸为4 096×4 096，数据分辨率为0.5-1m，涵盖不同季节、不同天气、不同视点、不同背景、不同尺度和不同质量的影像，具有丰富的表现形式，部分数据样本标签如图8所示，以旋转框的方式对飞机进行标注，影像背景比较复杂，云雾、周边建筑等其他干扰因素多，目标尺寸多样，部分小目标排列密集，一些飞机目标颜色与背景相近，不容易准确识别。“天智杯”数据集包含11类目标，由于部分型号在训练集和验证集中数量极少，本案可选取其中9类进行试验。部分目标样例如图9所示，飞机型号众多，单一类别飞机可能由1个或多个型号组成，因此同一类别飞机外观存在不一致现象，同时，不同类别的飞机之间存在外观相似性，比如E、F、H、I这4类目标存在形状类似的现象，类间特征差别较小。

[0103] 此外，DOTA数据集中的飞机目标和数据增广手段将使用于本文的级联网络中。DOTA数据集中197张大尺寸遥感影像融合“天智杯”数据集为定向检测网络提供充分的训练数据。针对分类网络，可制作用于训练机型分类的数据集，部分样本来自于“天智杯”训练样本数据，部分样本截取自“谷歌地球”，采用缩放、旋转、亮度变化、对比度变化、场景迁移等数据增广操作进行了目标样本制作和扩充，其中训练样本24 277个，测试样本6412个，数据增强效果如图10所示。

[0104] 实验运行环境为Ubuntu 20.04, GPU为NVIDIA Quadro RTX 5000(显存为16G), CPU为Intel (R) Xeon (R) Gold 6128, CUDA10.2+Pytorch1.7.1。对于定向检测网络，设置训练循环周期epochs=120, batch_size=4, 选择随机梯度下降法SGD作为优化器，设置动量因子为0.937, 权重衰减系数为0.000 5, 初始学习率设为0.05, 训练120轮，通过设置函数 $\text{lr_sc} = 0.05 * ((1 + \cos(x * \pi / 120)) / 2 + 0.01)$, $x \in [0, 119]$, 使学习率平滑地从0.05降至0.0005。对于分类网络，设置训练循环周期epochs=80, batch_size=16, 选择衰减系数为0.9、动量因

子为0.9的均方根传递算法 (RMSProp) 作为优化器, 权重衰减系数为0.000 05, 初始学习率设为0.01, 每2.4个epoch学习率下降至97%。

[0105] 首先, 通过多组实验研究纵横比权重函数 (ARSWF) 的有效性及其超参数 α 的选择。从DOTA数据集中筛选飞机 (PL)、储罐 (ST) 和环形交叉路口 (RA) 这三种类正方形的目标进行实验。实验结果如表2所示, mAP50表示IoU阈值取0.5的三类目标平均精度, 结果表明, 使用纵横比感知权重函数可以加速损失函数收敛, 促进类正方形物体检测精度的提高。 α 决定了损失抑制窗口的大小, 当 $\alpha=0$ 的时候, ARSWF是一个恒定的函数0, 完全抑制了角度损失值, 无法进行训练; 当 α 比较小的时候, 函数的抑制窗口相对较大, 它会抑制正常的角度损失值, 影响正常的训练效果。当 α 很大的时候, 抑制窗口接近脉冲函数, ARSWF在特定角度值附近的容错率会降低。实验结果也表明, α 适当取值的效果是最好的, 三类目标的总体精度提高了1.5%。

[0106] 表2在不同超参数 α 下的检测结果比较

	不使用	$\alpha=1$	$\alpha=5$	$\alpha=10$	$\alpha=20$	$\alpha=30$	$\alpha=40$
[0107] PL	94.4	91.8	94.2	94.8	95.2	95.1	95.0
ST	75.6	73.5	76.1	77.2	77.9	78.0	77.7
RA	58.8	54.6	59.0	59.4	60.3	59.9	59.7
mAP50	76.3	73.3	75.8	76.4	77.8	77.7	77.5

[0108] 为进一步验证旋转检测网络的性能, 在完整的DOTA测试集上进行了对比验证。表3结果显示, 与其他优秀的旋转网络相比, 本案算法具有良好的性能, 对于大纵横比目标, 如桥梁 (BD)、船舶 (SH) 和港口 (HA) 等, 仍然具有良好的检测精度。检测效果图如图11所示, 面对复杂的影像背景、密集的小目标、以及大纵横比的物体都获得了不错的检测效果。

[0109] 表1本案算法与其他优秀旋转检测算法在DOTA数据集上的对比结果

模型	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP50
RoI-Transformer	88.64	78.52	43.44	75.92	68.81	73.68	83.59	90.74	77.27	81.46	58.39	53.54	62.83	58.93	47.67	69.56
O2-DNet	89.31	82.14	47.33	61.21	71.32	74.03	78.62	90.76	82.23	81.36	60.93	60.17	58.21	66.98	61.03	71.04
P-RSDet	88.58	77.84	50.44	69.29	71.10	75.79	78.66	90.88	80.10	81.71	57.92	63.03	66.30	69.77	63.13	72.30
SCRDet	89.98	80.65	52.09	68.36	68.36	60.32	72.41	90.85	87.94	86.86	65.02	66.68	66.25	68.24	65.21	72.61
FADet	90.21	79.58	45.49	76.41	73.18	68.27	79.56	90.83	83.40	84.68	53.40	65.42	74.17	69.69	64.86	73.28
R3Det	89.49	81.17	50.53	66.10	70.92	78.66	78.21	90.81	85.26	84.23	61.81	63.77	68.16	69.83	67.17	73.74
Gliding Vertex	89.64	85.00	52.26	77.34	73.01	73.14	86.82	90.74	79.02	86.81	59.55	70.91	72.94	70.86	57.32	75.02
BBAVectors	88.63	84.06	52.13	69.56	78.26	80.40	88.06	90.87	87.23	86.39	56.11	65.62	67.10	72.08	63.96	75.36
FFA	90.1	82.7	54.2	75.2	71.0	79.9	83.5	90.7	83.9	84.6	61.2	68.0	70.7	76.0	63.7	75.7
FPN-CSL	90.25	85.53	54.64	75.31	70.44	73.51	77.62	90.84	86.15	86.69	69.60	68.04	73.83	71.10	68.93	76.17
本案算法	90.46	85.69	54.67	65.45	81.13	83.43	88.34	90.75	87.17	88.23	59.79	68.44	76.72	79.58	66.40	77.75

[0111] 其次, 为了测试改进分类网络对机型分类性能的提升效果, 在自建数据集上, 设计了多组消融实验, 如表4所示, 可以看出, 本案方案中两个改进模块都可以提升EfficientNetV2-S进行机型分类的效果, 集成学习模块提升了1.2%的分类精度, 但增加了1M参数量, 说明使用多个子分类器抽取不完全一样的特征, 可以更加充分地解析目标图像的信息, 从而获得比单一分类器更好的分类效果; Fisher正则化器提升了1个百分点, 在一定程度上可以减小类内离散度和扩大类间离散度, 而且没有增加参数负担, 这两个模块组合的分类精度最好, 共提升2%的分类精度。

[0112] 表4分类网络消融实验结果

	EfficientNetV2-S	集成学习	Fisher 正则化	Top1_Acc	参数量
	✓			91.8%	14M
[0113]	✓	✓		93.0%	15M
	✓		✓	92.8%	14M
	✓	✓	✓	93.8%	15M

[0114] 图12是改进分类网络在测试集上得到的混淆矩阵,对角线为各机型分类的准确率,其余位置为分类错误率,颜色越深代表占比越大。可以看到9类型号中有8类型号的准确率达到90%以上,I类型号的准确率最低为88.7%,其中A类错分成E类较多,C、F、H、I类目标相互错分较多。

[0115] 然后,算法由“YOLOv5+角度分类器(CSL)+分类网络”构成,为验证多网络级联的效果,首先在FDDA数据集上与端到端的单级网络“YOLOv5+CSL”进行对比,实验中,在级联网络里使用与YOLOv5+CSL相同的训练权重实现一类飞机目标的定向检测。图13是单级网络检测结果的混淆矩阵图,表5是级联网络与单级网络的实验结果对比。从图13中可以看出,除了存在背景误检和目标漏检的因素外,单级网络对机型分类的好坏很大程度上影响了检测精度,比如F-16和F-22被过多地划分为F-15,导致战斗机类目标检测精度整体偏低。表5实验结果表明,在飞机定位精度相同的前提下,使用级联网络可以实现更加精准的型号分类。

[0116] 表5在FDDA测试集上的实验结果对比

模型	B-52	KC-10	E-3	B-1B	F-16	C-130	F-15	C-17	KC-135	E-8	F-22	波音-737	直升机	mAP50
[0117] YOLOv5+CSL	95.5	96.8	97.6	97.8	78.1	95.4	88.8	94.7	96.3	88.4	75.2	93.0	90.5	91.4
级联网络	95.3	97.1	97.3	97.6	82.9	95.4	91.2	95.0	96.8	88.7	79.6	93.2	90.3	92.3

[0118] 表6级联网络与其他算法的检测性能对比

Method	A	B	C	D	E	F	G	H	I	mAP50
[0119] CSL	83.7	76.4	87.9	61.2	73.9	89.9	85.0	92.8	80.5	81.3
BBAVectors	75.4	70.5	79.4	54.9	78.7	86.6	77.5	85.9	70.2	75.5
R ³ Det	78.7	79.1	82.4	59.4	66.9	85.7	78.6	91.5	70.8	77.0
本案算法	87.6	78.2	87.0	64.6	77.2	91.3	89.1	94.2	84.3	83.7

[0120] 表6展示了级联网络与CSL、BBAVectors、R³Det等优秀旋转检测算法在“天智杯”数据集上的检测性能对比结果,图14是他们对各类目标检测精度的柱状图。根据表6和图14可知,相对于其他3种方法,本案算法在A、D、G、I这五类飞机目标中的检测精度得到了较大的提升。其中D、G是战斗机和MV-22等类型的小目标,在低分辨率和复杂背景中较难精准检测,本案中级联网络在一定程度上提升了小目标的检测率;A、E、F、H、I类目标外观特征比较相近,有一些机型改装自民航客机,差别较小,容易分类错误,本案中二级细粒度分类网络能够较好地区分它们之间的特征差别,显著提高了模型的性能,在这四类网络中表现最好。

[0121] 为了更直观地展现改进算法的检测性能,实验截取3个具有代表性的测试集场景,依次是光线暗淡、云雾缭绕的机场;分布具有与民航特征相似机型的机场;分布较多小目标的机场。如图15所示,从检测结果中看出,在基本能够检测出飞机目标的前提下,本案算法的分类效果是最好的,在3类场景中,仅有两架飞机分类错误和一架直升机漏检,而其他算法在区分F、H、I类飞机上面出现较多错误,CSL和BBVectors对小目标存在更多漏检。

[0122] 最后,为验证本案算法具有良好的泛化性能,从“谷歌地球”上获取某一未知机场

的遥感影像,尺寸大小为10 696*6 204,分辨率为1m,整体检测效果如图16所示。选择A、B、C三个区域放大分析,这三个区域的飞机目标均实现了全部检测,没有出现虚警和漏警。在机型分类效果上,仅有A区域1架和C区域1架飞机分类错误;而在目标方向检测中,由于D型号飞机偏小,像素信息不足,在检测的时候难以分辨排列方向,区域B有5架飞机的机头方向与检测框出现了偏差,这也是导致D类飞机检测精度较低的一个重要原因。但总体来说,本案中级联网络可以有效应对不同场景、不同尺寸,不同排列下的多型号飞机目标检测,具有良好的泛化性能。

[0123] 除非另外具体说明,否则在这些实施例中阐述的部件和步骤的相对步骤、数字表达式和数值并不限制本发明的范围。

[0124] 基于上述的方法和/或系统,本发明实施例还提供一种服务器,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上述的方法。

[0125] 基于上述的方法和/或系统,本发明实施例还提供一种计算机可读介质,其上存储有计算机程序,其中,该程序被处理器执行时实现上述的方法。

[0126] 在这里示出和描述的所有示例中,任何具体值应被解释为仅仅是示例性的,而不是作为限制,因此,示例性实施例的其他示例可以具有不同的值。

[0127] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步定义和解释。

[0128] 最后应说明的是:以上所述实施例,仅为本发明的具体实施方式,用以说明本发明的技术方案,而非对其限制,本发明的保护范围并不局限于此,尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,其依然可以对前述实施例所记载的技术方案进行修改或可轻易想到变化,或者对其中部分技术特征进行等同替换;而这些修改、变化或者替换,并不使相应技术方案的本质脱离本发明实施例技术方案的精神和范围,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应所述以权利要求的保护范围为准。

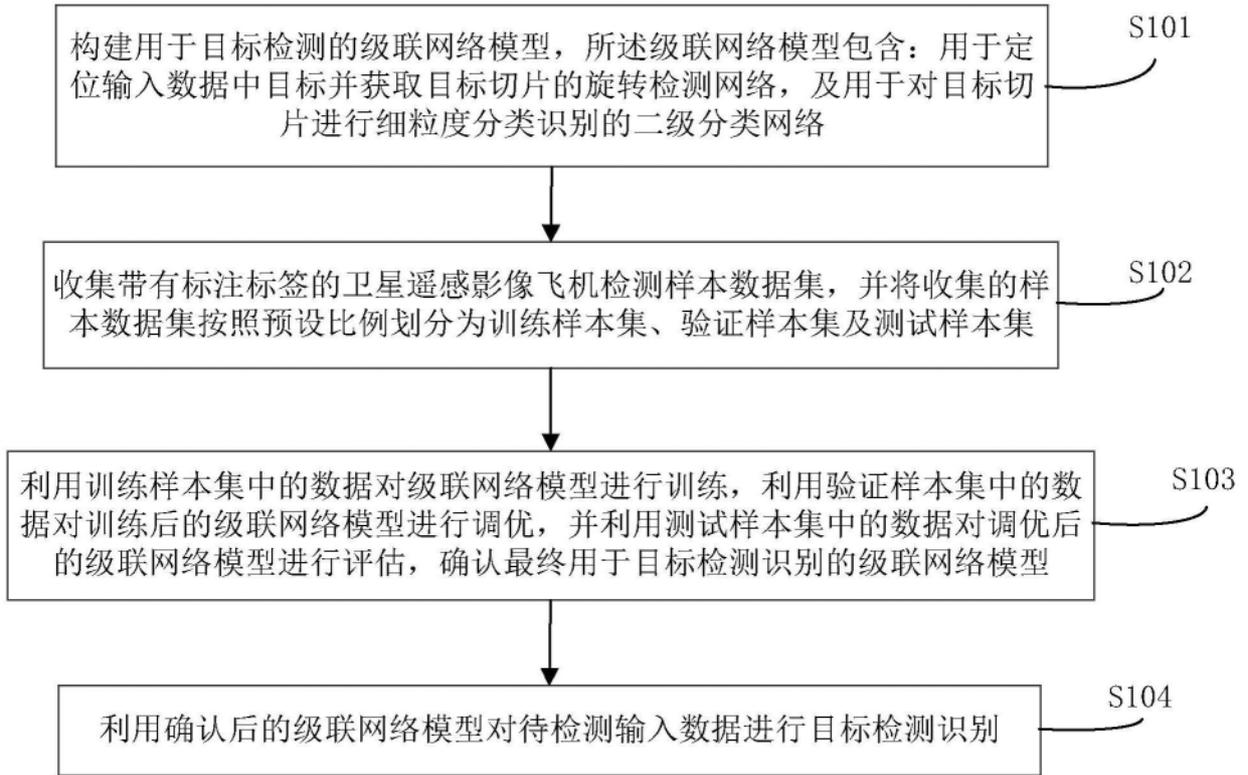


图1

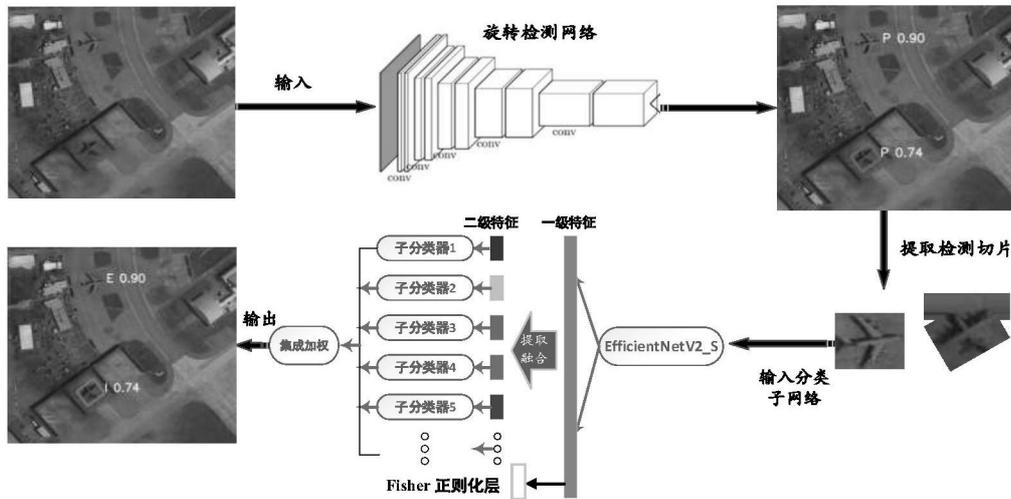


图2

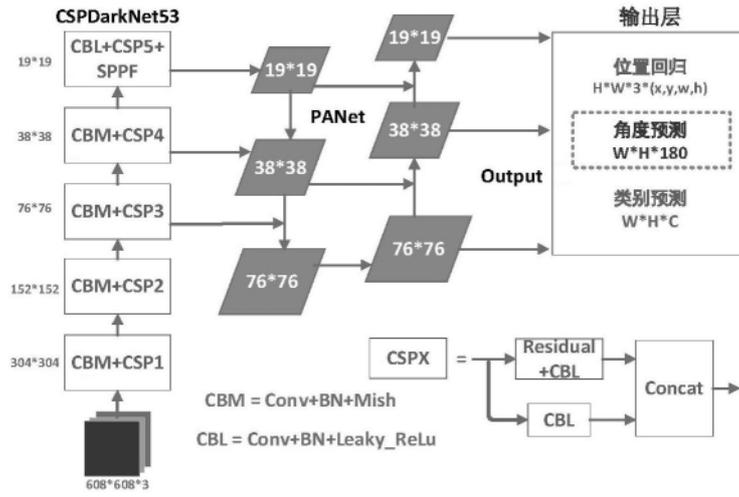


图3

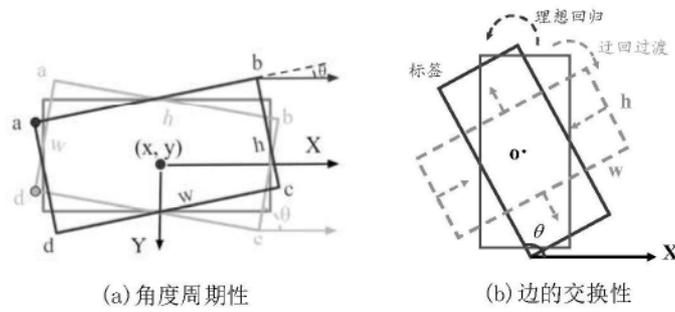


图4

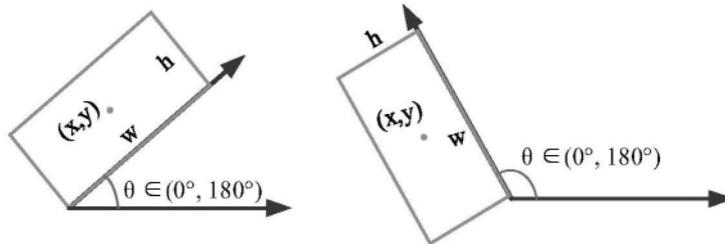


图5

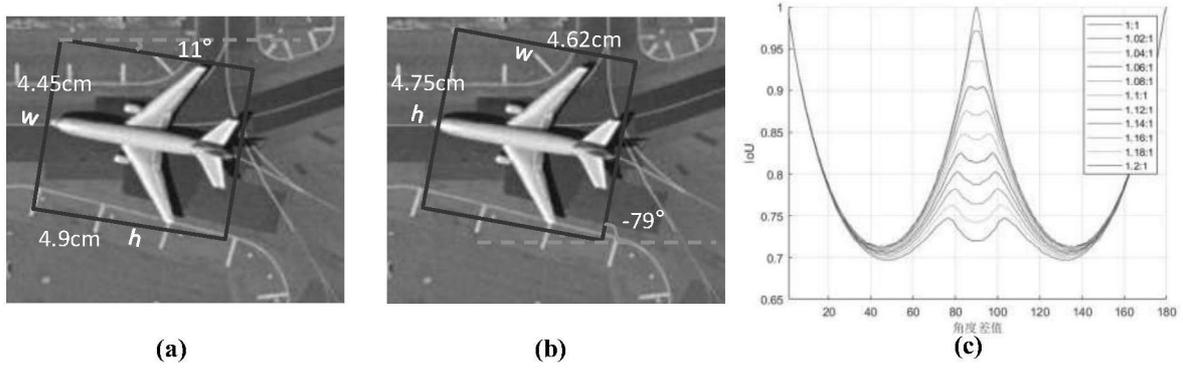


图6

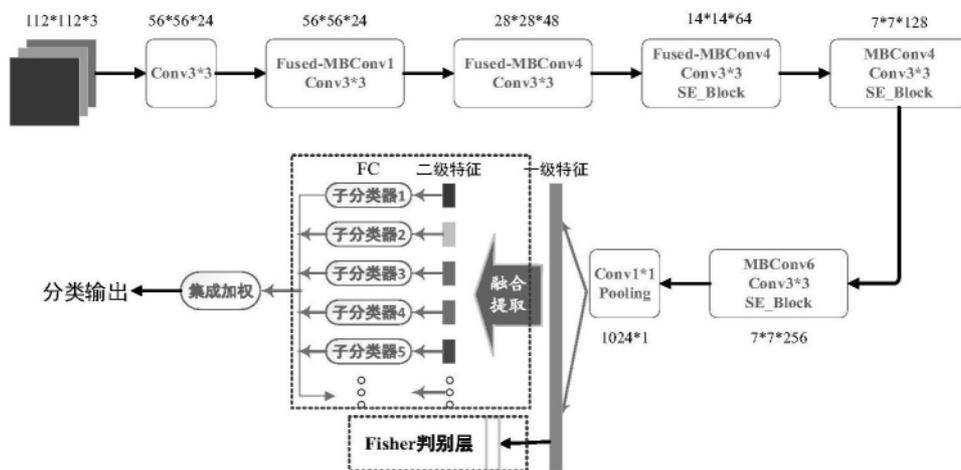


图7



图8

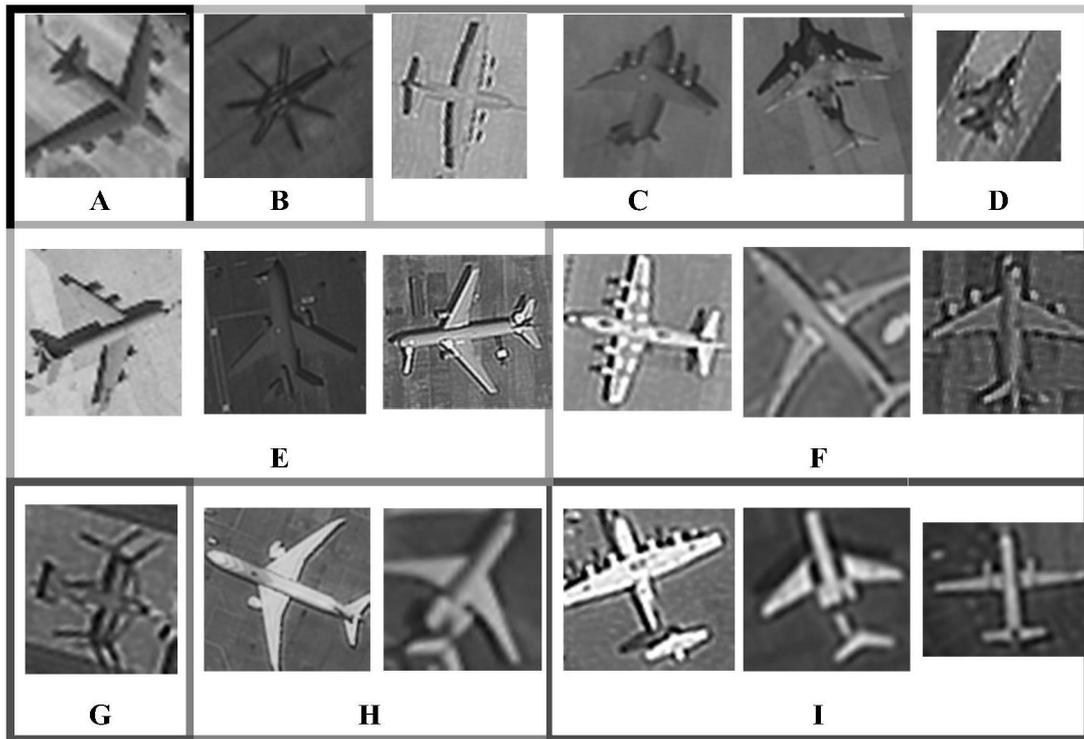


图9

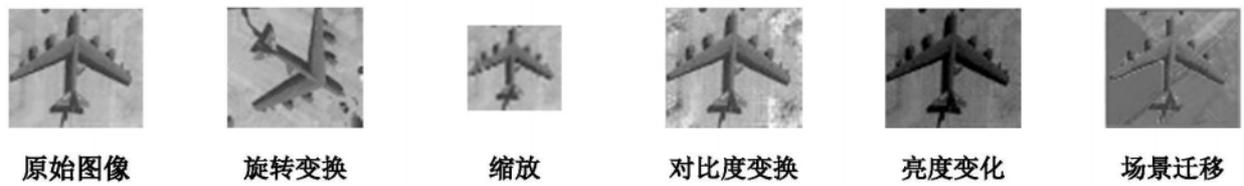


图10



图11

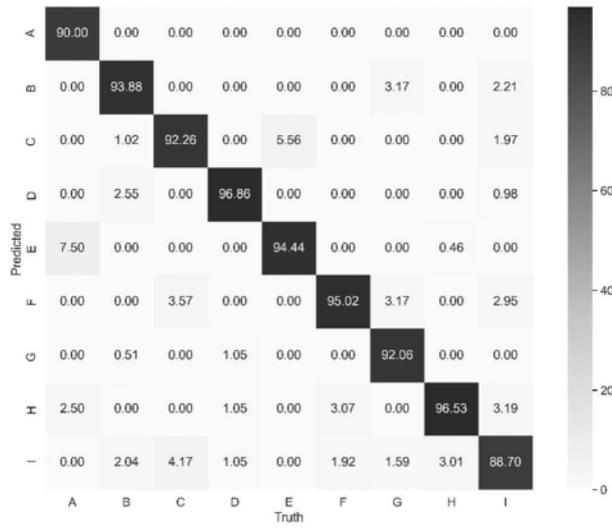


图12

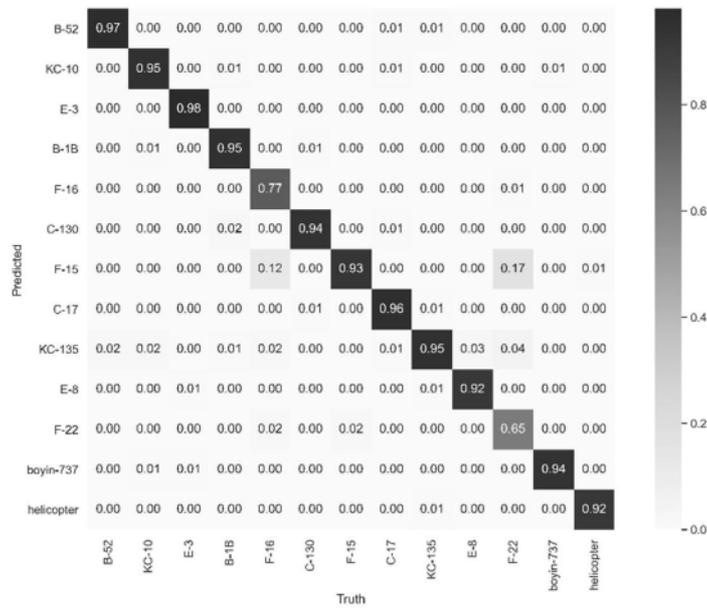


图13

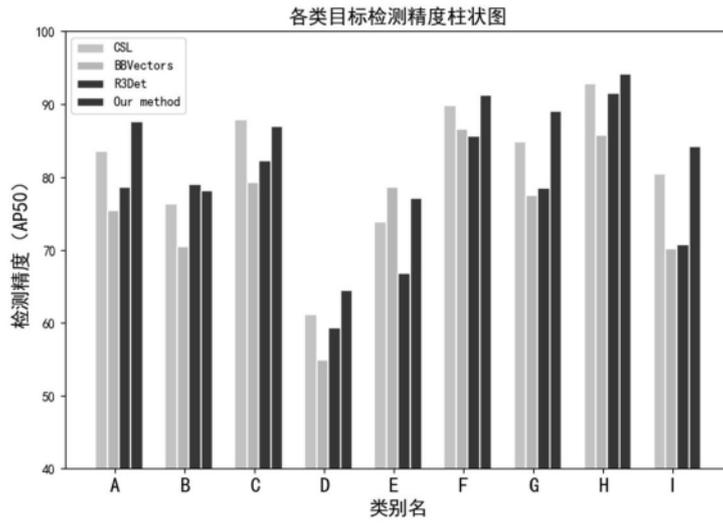
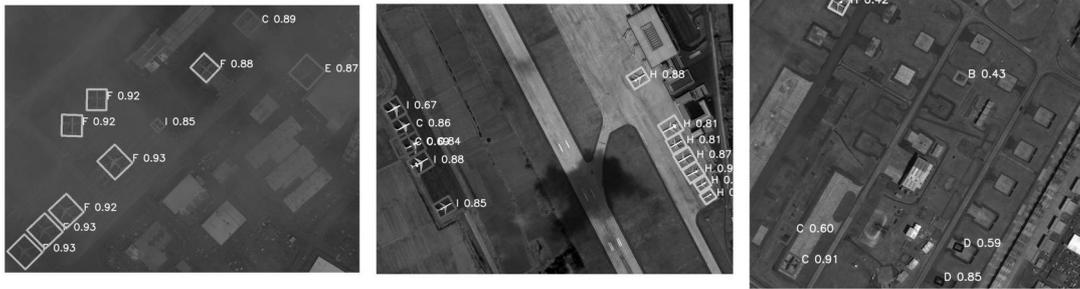
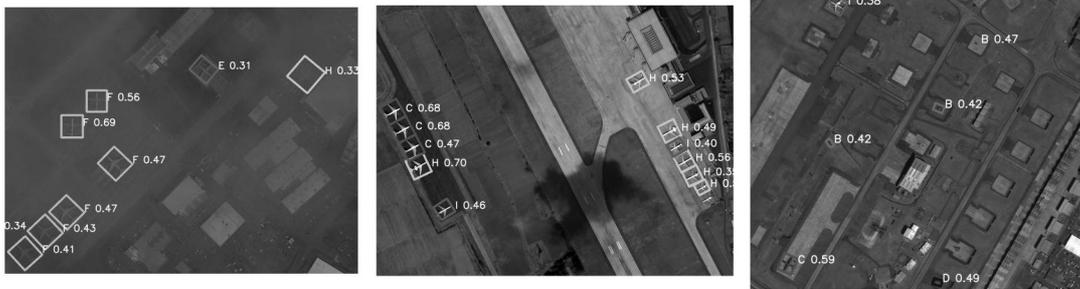


图14

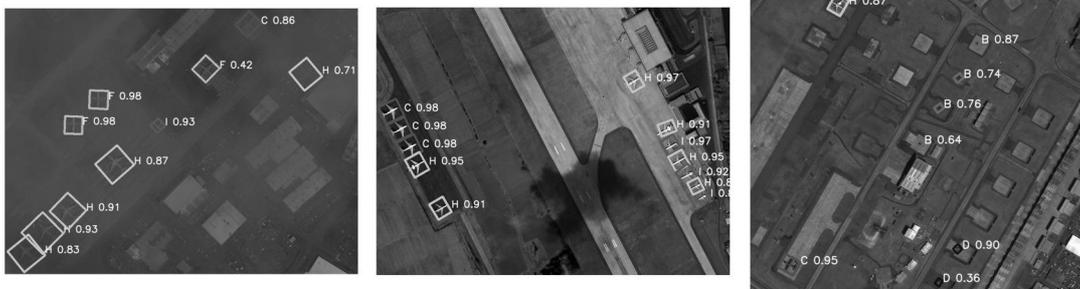




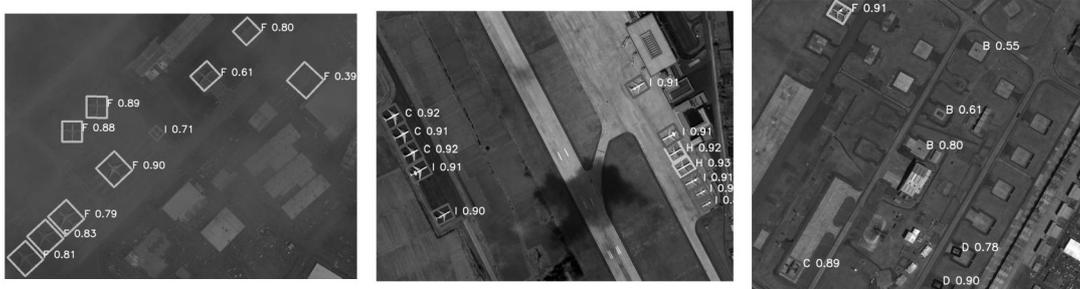
(b) CSL 检测结果



(c) BBVectors 检测结果



(d) R³Det 检测结果



(e) 本案算法检测结果

图15



图16