(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 114677554 A
(43)申请公布日 2022.06.28

(21)申请号 202210179351.8

(22)申请日 2022.02.25

(71)申请人 华东理工大学
地址 200237 上海市徐汇区梅陇路130号

(72)发明人 杨文 舒浩宇 陈培陪 顾凯峰 刘奉奉 田腾飞

(74)专利代理机构 上海智信专利代理有限公司 31002
专利代理师 杨怡清

(51)Int.Cl.
G06V 10/774(2022.01)
G06V 10/143(2022.01)
G06V 10/74(2022.01)
G06V 10/50(2022.01)
G06K 9/62(2022.01)

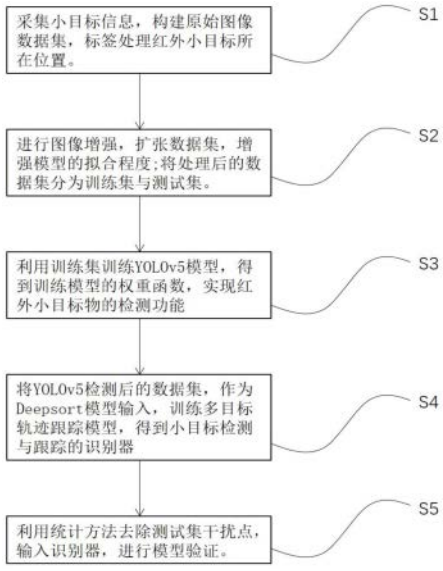
权利要求书2页 说明书9页 附图10页

(54)发明名称

一种基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法

(57)摘要

本发明涉及一种基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法，包括：步骤S1，利用红外成像设备采集复杂背景下的红外小目标图像，获取红外小目标图像数据集；步骤S2，对红外小目标图像数据集进行预处理，获取预处理后的图像数据集，并将预处理后的图像数据集划分为训练集和验证集；步骤S3，对YOLOv5算法中的YOLOv5s模型进行训练，获取Deepsort模型的训练集；步骤S4，将Deepsort模型的训练集输入至Deepsort模型进行训练，构建红外小目标检测与跟踪的识别器；步骤S5，利用红外小目标检测与跟踪的识别器，对红外小目标进行实时检测跟踪。本发明能够准确且快速地将复杂背景下的红外小目标检测出来，提高鲁棒性和检测率。并且，本发明可以达到实时跟踪效果。



CN 114677554 A

1.一种基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,包括:

步骤S1,利用红外成像设备采集复杂背景下的红外小目标图像,获取红外小目标图像数据集;

步骤S2,对所述红外小目标图像数据集进行预处理,获取预处理后的图像数据集,并将预处理后的图像数据集划分为训练集和验证集;

步骤S3,根据所述训练集和所述验证集,对YOLOv5算法中的YOLOv5s模型进行训练,获取Deepsort模型的训练集;

步骤S4,将所述Deepsort模型的训练集输入至Deepsort模型进行训练,获取Deepsort模型的权重参数,构建红外小目标检测与跟踪的识别器;

步骤S5,利用所述红外小目标检测与跟踪的识别器,对红外小目标进行实时检测跟踪。

2.根据权利要求1所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所述步骤S1包括:

步骤S11,使用红外成像设备在不同的复杂背景下对一个或多个飞行器进行连续帧拍摄采样,获取初始图像集;

步骤S12,在所述初始图像集的每个图像中标注出真实目标检测框,构建红外小目标的标签信息,初始图像集以及红外小目标的标签信息共同构成红外小目标图像数据集。

3.根据权利要求1所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所述步骤S2中对所述红外小目标图像数据集进行预处理的方法包括:

步骤S21,对所述红外小目标图像数据集进行数据清洗,得到清洗后图像数据集;

步骤S22,从所述清洗后图像数据集中随机选取若干帧图像,对选取出的若干帧图像进行数据增强,得到增强图像数据集;

步骤S23,从所述清洗后图像数据集中提取出每帧图像的背景图像,对提取出的背景图像进行混合处理,得到混合背景图像集;

步骤S24,从所述红外小目标图像数据集中提取出每帧图像所含噪点,将提取出的噪点随机粘贴至所述清洗后图像数据集、所述增强图像数据集和所述混合背景图像集中;

步骤S25,将所述步骤S24中随机粘贴噪点的清洗后图像数据集、增强图像数据集和混合背景图像集进行合并,得到最终的预处理后的图像数据集。

4.根据权利要求3所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所述步骤22中数据增强的方法为粘贴增强法或Mosaic增强法。

5.根据权利要求1所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所述步骤S3包括:

步骤S31,初始化YOLOv5s模型参数,包括:批次处理量、迭代次数、图像分辨率、交并比阈值以及置信度阈值;

步骤S32,将所述训练集输入至YOLOv5s模型,根据初始化的YOLOv5s模型参数,输出每次迭代后的预测目标检测框;

步骤S33,将每次迭代后的预测目标检测框与标签信息中的真实目标检测框进行IOU损失计算,得到每次迭代后的学习权重;

步骤S34,针对每次迭代后的学习权重,根据所述验证集,选取在验证集上测试误差最

小的学习权重,将该测试误差最小的学习权重作为YOLOv5s模型的权重参数;

步骤S35,根据YOLOv5s模型的权重参数,获取每帧图像中的目标识别候选框,每帧图像中的目标识别候选框构成Deepsort模型的训练集。

6.根据权利要求5所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所述批次处理量设为64,所述迭代次数设为2000,所述图像分辨率设为640*640,所述交并比阈值设为0.4,所述置信度阈值设为0.6。

7.根据权利要求1所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所述步骤S4包括:

步骤S41,根据Deepsort模型的训练集,确定红外小目标在每帧图像中的真实位置,并根据红外小目标在第k-1(k≥2)帧图像中的真实位置,采用卡尔曼滤波来确定红外小目标在第k帧图像中的预测位置;

步骤S42,采用匈牙利算法对红外小目标在第k帧图像中的预测位置和红外小目标在第k帧图像中的真实位置进行级联匹配,得到初次匹配成功的结果、初次未匹配的轨迹以及初次未匹配的检测框;

步骤S43,对步骤S42中初次未匹配的轨迹和初次未匹配的检测框进行IOU匹配,得到再次匹配成功的结果、再次未匹配的轨迹以及再次未匹配的检测框;

步骤S44,根据初次匹配成功的结果和再次匹配成功的结果,更新卡尔曼滤波的参数;

步骤S45,为再次未匹配的检测框分配新的轨迹和新的ID,并通过ReID提取检测框中目标物的特征集;同时判断再次未匹配的轨迹是否处于确定状态,保留处于确定状态且失配数小于30帧的轨迹,重复步骤S41-步骤S44。

8.根据权利要求1所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所步骤S5包括:

步骤S51,设定红外小目标检测与跟踪的识别器中的参数,并将YOLOv5s模型的权重参数加载至YOLOv5s模型,将Deepsort模型的权重参数加载至Deepsort模型;

步骤S52,利用红外成像设备采集红外小目标的实时图像,并对采集到的实时图像进行对比度增强处理,获取增强后的实时图像;

步骤S53,对增强后的实时图像的前100帧进行像素平均统计,获取实时图像中的盲点,将每个盲点的位置存入盲点列表;

步骤S54,将增强后的实时图像输入YOLOV5s模型,生成置信度在0.01以上的实时目标候选框集;

步骤S55,将实时目标候选框集输入至Deepsort模型,获取滤波后的实时目标候选框集,并剔除滤波后的实时目标候选框集中包含盲点列表的目标候选框,生成最终的目标检测框,实现红外小目标的实时跟踪。

9.跟据权利要求7所述的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,其特征在于,所述步骤S52中对采集到的实时图像进行对比度增强处理的方法为:获取每一帧图像的直方图分布,将每帧图像的直方图分布改变成近似均匀分布直方图。

一种基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法

技术领域

[0001] 本发明涉及目标检测与跟踪领域，更具体地涉及一种基于YOLOv5与 Deepsort的统计滤波红外小目标检测跟踪方法。

背景技术

[0002] 目前，目标检测技术在各个领域均具有广泛运用。而红外小目标检测一直以来是红外图像处理领域中的热点，其技术研究对军事预警、模式识别、图像处理等领域具有重要意义。然而，在实际场景下，红外图像中的目标成像距离较长，通常以点状的形式出现。并且红外信号在传播过程中受空气削弱严重，导致其通常以高斯分布的点目标形式存在，同时其自身不具有显著形状和纹理信息。另外，在复杂背景下的红外小目标常被噪声和杂波淹没，导致含有小目标的红外图像信噪比很低。这些特点给红外小目标检测技术带来了极大挑战。

[0003] 传统目标检测算法包括区域框选择、特征提取、归分类别，其存在两个问题：一是区域选择策略没有针对性，时间复杂度高；二是鲁棒性差。而随着深度学习在计算机视觉领域的不断发展，其在目标检测方面得到了重大突破，目前主要分为两类：一类为基于检测框和分类器的二阶算法，如R-CNN (区域卷积神经网络)系列，这类算法精度高，但是网络结构复杂导致检测速度慢，难以满足实时的目标检测；另一类为基于回归的一阶算法，如YOLO(你只看一次)系列。这类算法推理速度快，可以满足实时检测。YOLO系列的第5代版本YOLOv5拥有更轻量的框架以及更快的推理速度，但其依然只能对目标物进行识别检测，无法进行轨迹跟踪。并且，YOLOv5在环境条件不好的情况下，十分容易发生漏检。

[0004] Deepsort(深度排序)跟踪算法利用卡尔曼滤波对目标进行预测和更新，并在级联匹配中通过匈牙利算法将目标轨迹的预测框和检测框进行数据关联匹配，能良好跟踪目标。但现有的结合YOLOv5与Deepsort跟踪算法的检测跟踪方法所针对的对象均为大目标物，例如行人。因而需要研发一种针对复杂背景下的红外小目标的检测跟踪方法。

发明内容

[0005] 为解决上述现有技术中的问题，本发明提供一种基于YOLOv5与 Deepsort的统计滤波红外小目标检测跟踪方法，能够准确、快速地实现背景分离，检测并跟踪在复杂背景下的红外小目标，同时能够提高鲁棒性和检测率。

[0006] 本发明提供的一种基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法，包括：

[0007] 步骤S1，利用红外成像设备采集复杂背景下的红外小目标图像，获取红外小目标图像数据集；

[0008] 步骤S2，对所述红外小目标图像数据集进行预处理，获取预处理后的图像数据集，并将预处理后的图像数据集划分为训练集和验证集；

[0009] 步骤S3,根据所述训练集和所述验证集,对YOLOv5算法中的 YOLOv5s模型进行训练,获取Deepsort模型的训练集;

[0010] 步骤S4,将所述Deepsort模型的训练集输入至Deepsort模型进行训练,获取Deepsort模型的权重参数,构建红外小目标检测与跟踪的识别器;

[0011] 步骤S5,利用所述红外小目标检测与跟踪的识别器,对红外小目标进行实时检测跟踪。

[0012] 进一步地,所述步骤S1包括:

[0013] 步骤S11,使用红外成像设备在不同的复杂背景下对一个或多个飞行器进行连续帧拍摄采样,获取初始图像集;

[0014] 步骤S12,在所述初始图像集的每个图像中标注出真实目标检测框,构建红外小目标的标签信息,初始图像集以及红外小目标的标签信息共同构成红外小目标图像数据集。

[0015] 进一步地,所述步骤S2中对所述红外小目标图像数据集进行预处理的方法包括:

[0016] 步骤S21,对所述红外小目标图像数据集进行数据清洗,得到清洗后图像数据集;

[0017] 步骤S22,从所述清洗后图像数据集中随机选取若干帧图像,对选取出的若干帧图像进行数据增强,得到增强图像数据集;

[0018] 步骤S23,从所述清洗后图像数据集中提取出每帧图像的背景图像,对提取出的背景图像进行混合处理,得到混合背景图像集;

[0019] 步骤S24,从所述红外小目标图像数据集中提取出每帧图像所含噪点,将提取出的噪点随机粘贴至所述清洗后图像数据集、所述增强图像数据集和所述混合背景图像集中;

[0020] 步骤S25,将所述步骤S24中随机粘贴噪点的清洗后图像数据集、增强图像数据集和混合背景图像集进行合并,得到最终的预处理后的图像数据集。

[0021] 进一步地,所述步骤22中数据增强的方法为粘贴增强法或Mosaic增强法。

[0022] 进一步地,所述步骤S3包括:

[0023] 步骤S31,初始化YOLOv5s模型参数,包括:批次处理量、迭代次数、图像分辨率、交并比阈值以及置信度阈值;

[0024] 步骤S32,将所述训练集输入至YOLOv5s模型,根据初始化的YOLOv5s 模型参数,输出每次迭代后的预测目标检测框;

[0025] 步骤S33,将每次迭代后的预测目标检测框与标签信息中的真实目标检测框进行IOU损失计算,得到每次迭代后的学习权重;

[0026] 步骤S34,针对每次迭代后的学习权重,根据所述验证集,选取在验证集上测试误差最小的学习权重,将该测试误差最小的学习权重作为YOLOv5s 模型的权重参数;

[0027] 步骤S35,根据YOLOv5s模型的权重参数,获取每帧图像中的目标识别候选框,每帧图像中的目标识别候选框构成Deepsort模型的训练集。

[0028] 进一步地,所述批次处理量设为64,所述迭代次数设为2000,所述图像分辨率设为640*640,所述交并比阈值设为0.4,所述置信度阈值设为0.6。

[0029] 进一步地,所述步骤S4包括:

[0030] 步骤S41,根据Deepsort模型的训练集,确定红外小目标在每帧图像中的真实位置,并根据红外小目标在第k-1(k≥2)帧图像中的真实位置,采用卡尔曼滤波来确定红外小目标在第k帧图像中的预测位置;

[0031] 步骤S42,采用匈牙利算法对红外小目标在第k帧图像中的预测位置和红外小目标在第k帧图像中的真实位置进行级联匹配,得到初次匹配成功的结果、初次未匹配的轨迹以及初次未匹配的检测框;

[0032] 步骤S43,对步骤S42中初次未匹配的轨迹和初次未匹配的检测框进行 IOU匹配,得到再次匹配成功的结果、再次未匹配的轨迹以及再次未匹配的检测框;

[0033] 步骤S44,根据初次匹配成功的结果和再次匹配成功的结果,更新卡尔曼滤波的参数;

[0034] 步骤S45,为再次未匹配的检测框分配新的轨迹和新的ID,并通过ReID 提取检测框中目标物的特征集;同时判断再次未匹配的轨迹是否处于确定状态,保留处于确定状态且失配数小于30帧的轨迹,重复步骤S41-步骤S44。

[0035] ,所步骤S5包括:

[0036] 步骤S51,设定红外小目标检测与跟踪的识别器中的参数,并将 YOLOv5s模型的权重参数加载至YOLOv5s模型,将Deepsort模型的权重参数加载至Deepsort模型;

[0037] 步骤S52,利用红外成像设备采集红外小目标的实时图像,并对采集到的实时图像进行对比度增强处理,获取增强后的实时图像;

[0038] 步骤S53,对增强后的实时图像的前100帧进行像素平均统计,获取实时图像中的盲点,将每个盲点的位置存入盲点列表;

[0039] 步骤S54,将增强后的实时图像输入YOLOV5s模型,生成置信度在 0.01以上的实时目标候选框集;

[0040] 步骤S55,将实时目标候选框集输入至Deepsort模型,获取滤波后的实时目标候选框集,并剔除滤波后的实时目标候选框集中包含盲点列表的目标候选框,生成最终的目标检测框,实现红外小目标的实时跟踪。

[0041] 进一步地,所述步骤S52中对采集到的实时图像进行对比度增强处理的方法为:获取每一帧图像的直方图分布,将每帧图像的直方图分布改变成近似均匀分布直方图。

[0042] 本发明结合用于目标检测的YOLOv5框架和用于目标跟踪的Deepsort 框架,并针对红外小目标的复杂背景设计了数据清洗及增强等预处理和背景去除的后处理,可以克服背景中的干扰点和盲元点,从而能够准确且快速地将复杂背景下的红外小目标检测出来,提高鲁棒性和检测率。并且,本发明可以达到实时跟踪效果。

附图说明

[0043] 图1是按照本发明的基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法的流程图。

[0044] 图2是标注出的红外小目标的标签信息示意图。

[0045] 图3是图1中步骤S4的流程图。

[0046] 图4(a)-图4(h)是置信度阈值为0.001下的上述各性能指标图。

[0047] 图5(a)是本发明识别检测模型测试集上的P-R曲线图;图5(b)是不同置信度阈值下,F1分数的变化图;图5(c)是不同置信度阈值下,对应的精确率变化图;图5(d)是不同置信度阈值下,对应的召回率变化图。

[0048] 图6(a)-图6(c)是最终跟踪预测效果图。

具体实施方式

[0049] 下面结合附图,给出本发明的较佳实施例,并予以详细描述。

[0050] 如图1所示,本发明提供一种基于YOLOv5与Deepsort的统计滤波红外小目标检测跟踪方法,包括以下步骤:

[0051] 步骤S1,利用红外成像设备采集复杂背景下的红外小目标图像,获取红外小目标图像数据集。其中,复杂背景指的是树、少云天空、建筑物、连续变化天气、复杂云层、海面与海空等背景。

[0052] 具体地,步骤S1包括:

[0053] 步骤S11,使用红外成像设备在不同的复杂背景下对一个或多个飞行器进行连续帧拍摄采样,获取初始图像集。采样帧率可以为50ms、100ms或其他合适的帧率。在红外成像设备的拍摄过程中,由于传感器延迟、噪声干扰等影响,会采集到模糊图像或存在缺失的图像。针对这些模糊图像或存在缺失的图像,需要进行重采样处理,以获取清晰完整的初始图像集。

[0054] 步骤S12,在初始图像集的每个图像中标注出真实目标检测框,构建红外小目标的标签信息,初始图像集以及红外小目标的标签信息共同构成红外小目标图像数据集。

[0055] 其中,标注真实目标检测框的方法为人工标注,例如采用labeling工具进行标注,获取目标检测框左上角点的坐标(x,y)、目标检测框的长度h以及目标检测框的宽度w,如图2所示。前述标注出的x、y、h以及w即构成红外小目标的标签信息。

[0056] 由于人工标注会存在人为误差因素,例如因操作失误而严重偏离真实目标或者存在细微误差。因而,为了尽可能减小人工行为的误差,增强数据集信息强度,提高模型训练精度,需要使用可视化操作来判断标注是否正确。判断过程具体如下:将labeling生成的xml文件格式转换为txt文件格式,并将txt文件中的数据信息存储至一列表;利用opencv工具读取初始图像集中的图像信息;将列表存储的数据信息与读取出的图像信息进行对比,判断目标检测框是否与红外小目标所在区域良好重叠,从而获取标签信息的正确率。当目标检测框与红外小目标所在区域的重叠面积占据两区域总面积的90%以上时,目标检测框与红外小目标所在区域良好重叠,重叠面积在总面积中的占比即为标签信息的正确率。若标签信息的正确率低于90%,则需要人工重新设定标注信息,并重新计算正确率。

[0057] 步骤S2,对红外小目标图像数据集进行预处理,获取预处理后的图像数据集,并将预处理后的图像数据集划分为训练集和验证集。

[0058] 对红外小目标图像数据集进行预处理的方法包括:

[0059] 步骤S21,对红外小目标图像数据集进行数据清洗,得到清洗后图像数据集。

[0060] 数据清洗是指利用统计学的加权平均法或机器学习法来对缺省值和异常值进行处理。在图像采集过程中,可能会因为噪声干扰导致部分图像异常与缺失,因而需要进行清洗处理。在本实施例中,采用统计学的加权平均法进行数据清洗:检测红外小目标图像数据集中是否存在某帧图像缺失或某帧图像异常,若存在某帧图像缺失,则利用该帧图像的前后帧图像进行平均填补;若存在某帧图像异常(即与其前后两帧图像产生不连续的关联),则将该帧图像删除,并利用该帧图像的前后两帧图像进行平均填补。

[0061] 步骤S22,从清洗后图像数据集中随机选取若干帧图像,对选取出的若干帧图像进行数据增强,得到增强图像数据集。

[0062] 数据增强是指将原始数据进行随机翻转、放缩、剪切、旋转、粘贴、拼凑等操作扩张出额外的数据集。在本实施例中，采用粘贴增强法或Mosaic 增强法对选取出的若干帧图像进行数据增强。粘贴增强方法为：通过将图像进行比例为0.1～2倍的大规模尺度缩放后，再随机将一帧图像中的红外小目标粘贴到另一帧图像中，以此增加红外小目标的数量，从红外小目标数量角度来增强检测。Mosaic增强方法为：随机采用4帧图片，以随机缩放、裁剪、排序的方式进行拼接，使红外小目标的分布较为均匀，并使一帧图像中有多个红外小目标，从红外小目标分布及数量角度来增强检测。

[0063] 步骤S23，从清洗后图像数据集中提取出每帧图像的背景图像，对提取出的背景图像进行混合处理，得到混合背景图像集。得到混合背景的图像集，可以增强模型对复杂背景的识别能力，减小将背景物当作是目标物的误检率。提取背景图像的方法为：采用图像膨胀腐蚀的方法，将目标物和噪点腐蚀后，保留的图像即为背景图像。

[0064] 步骤S24，从红外小目标图像数据集中提取出每帧图像所含噪点，将提取出的噪点随机粘贴至上述清洗后图像数据集、增强图像数据集和混合背景图像集中以扩大噪点的数据分布。其中，噪点采用人为手段进行剪切提取。

[0065] 步骤S25，将步骤S24中随机粘贴噪点的清洗后图像数据集、增强图像数据集和混合背景图像集进行合并，得到最终的预处理后的图像数据集。

[0066] 得到预处理后的图像数据集之后，按照随机采样的方式将预处理后的图像数据集以8:2的比例划分为训练集和验证集。即0.8份为训练集，0.2份

为验证集。

[0067] 步骤S3，根据步骤S2划分出的训练集和验证集，对YOLOv5算法中的YOLOv5s模型进行训练，获取Deepsort模型的训练集。本发明采用 YOLOv5s模型是为了更快速地实现检测效果。

[0068] YOLOv5s模型可视为一个二阶段的目标检测算法，包括以ResNet为基础的基准网络层、用于输出目标检测结果的采用FPN+PAN结构的Neck网络层、输出层、以及经过非极大值抑制处理后的输出端。其中，基准网络层输出特征映射矩阵。Neck网络层采用FPN+PAN的结构来提升特征的多样性和鲁棒性，加强网络特征的融合能力。其中，FPN表示特征金字塔网络，利用自上而下的上采样提取图像的强语义特征(即目标物的形状)；PAN表示像素聚合网络，利用自下而上的网络提取图像的强定位特征(即目标物的位置)。FPN与PAN进行融合，能够实现形状与位置特征的聚合。输出层采用 GIoU_Loss作为Bounding box的损失函数，输出目标检测结果。得到目标检测结果后，进行后处理，采用非极大值抑制，以消除同一个目标上的多个框以及堆叠在一起的输出边界框。

[0069] 具体地，步骤S3包括：

[0070] 步骤S31，初始化YOLOv5s模型参数，包括：批次处理量(batchsize)、迭代次数(epoch)、图像分辨率、交并比(IOU)阈值以及置信度阈值。

[0071] 由于图像集较大，模型训练时不能直接将所有图像一同输入，否则会导致维度爆炸，因而采用分批次的方法，每次输出64帧图像。YOLOv5s模型网络包括前向传播与反向传播两个过程，当所有图像都经历过一次前向和反向传播后，称为迭代1次，迭代是为了进行梯度优化，找到使损失函数收敛时的权重。在本实施例中，batchsize设为64，epoch设为2000，图像分辨率设为640*640，交并比阈值设为0.4，置信度阈值设为0.6。

[0072] 步骤S32,将步骤S2划分出的训练集输入至YOLOv5s模型,根据初始化的YOLOv5s模型参数,输出每次迭代后的预测目标检测框。在YOLOv5s 模型的输出端输出目标检测框的过程中,需根据置信度对检测出的目标检测框进行过滤,删除置信度小于设定置信度阈值0.6的目标检测框,保留下来的目标检测框即为预测目标检测框。

[0073] 步骤S33,将每次迭代后的预测目标检测框与标签信息中的真实目标检测框进行IOU损失计算,得到每次迭代后的学习权重。

[0074] 步骤S34,针对每次迭代后的学习权重,根据步骤S2划分出的验证集,选取在验证集上测试误差最小(即精度最高)的学习权重,将该测试误差最小对应的学习权重作为YOLOv5s模型的权重参数。测试误差的获取方法具体为:将验证集输入至以当前次迭代得到的学习权重作为权重参数的YOLOv5s 模型中,得到包括标注信息的预测框以及表示是否存在目标物的置信度,通过计算IOU与分类的交叉熵的损失,从而得出测试误差。

[0075] 步骤S35,根据YOLOv5s模型的权重参数,进行目标物检测,获取每帧图像中的目标识别候选框,每帧图像中的目标识别候选框构成Deepsort模型的训练集。具体地,将步骤S2划分出的训练集重新输入至确定好权重参数的YOLOv5s模型,并根据设定好的置信度阈值得出目标识别候选框。例如,若存在两个目标物,则对应的真实目标框也应为两个,但由于前述对图像数据集进行了增强,因而通过YOLOv5s模型输出的锚框可能有10个、15个或其他数量,这些数量较多的锚框即为目标识别候选框。但这些锚框数量大于2,说明存在误检,也可能存在漏检,因而需将这些锚框作为Deepsort模型的训练集,输入至Deepsort模型进行下一步处理。

[0076] 即进行步骤S4,将Deepsort模型的训练集输入至Deepsort模型进行训练,获取Deepsort模型的权重参数,构建红外小目标检测与跟踪的识别器。

[0077] 需要说明的是,将Deepsort模型的训练集输入至Deepsort模型进行训练之前,需要将Deepsort模型的网络结构修改为能够处理像素宽度在1～20左右的小目标图像的网络结构,或者将目标识别候选框预处理至Deepsort模型能够正确辨别的大小。

[0078] 采用Deepsort模型进行训练目的在于滤除误检框,补充漏检框。具体地,如图3所示,步骤S4包括:

[0079] 步骤S41,根据Deepsort模型的训练集,确定红外小目标在每帧图像中的真实位置,并根据红外小目标在第k-1(k≥2)帧图像中的真实位置,采用卡尔曼滤波来确定红外小目标在第k帧图像中的预测位置。需要说明的是,该步骤中所指的真实位置为每帧图像中的目标识别候选框所框出的位置,不同于步骤S12中真实目标检测框所框出的位置。该真实位置是一个确定值,用于后续采用卡尔曼滤波时区分确定值和预测值。

[0080] 卡尔曼滤波用于预测轨迹,并根据预测的轨迹信息来滤除误检框。例如,若下一时刻的预测状态信息是[2,3,4,5],那么候选框内的[10,11,12,13] 检测信息即为误检框,需要将这种检测框滤除。也就是说,先根据置信度阈值做一次粗略的滤波,再根据卡尔曼滤波做一次精细的滤波。

[0081] 卡尔曼滤波主要包括锚框的八个系统状态:X＝(x,y,,r,,,r,), x和y分别表示锚框的中心点在图像中的横坐标和纵坐标,h表示锚框的高度, r表示锚框高度与宽度的比值;x,y,r,分别为当前状态下x,y,,的导数,表示目标物在上一帧图像中的状态信息。卡尔曼滤波算法就是根据当前的状态信息,预测下一时刻的状态信息。卡尔曼滤波公式为:

[0082] $x' = Ax$

[0083] $P' = APA^T + Q$

[0084] $y = z - Hx'$

[0085] $K = P'H^T(HP'H^T + R)^{-1}$

[0086] $P = (I - KH)P'$

[0087] 式中，A为系统矩阵，$A = \begin{bmatrix} 1 & 0 & 0 & 0 & dt & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & dt & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & dt & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & dt \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$，P是协方差矩阵，用于衡量预测值与真实值的误差；Q是状态发生转移时带有的过程噪声；K是增益矩阵，用于最小化P；R是对状态进行测量时存在测量噪声，一般为高斯白噪声；I是单位矩阵，H是测量矩阵。

[0088] 步骤S42，采用匈牙利算法对红外小目标在第k帧图像中的预测位置和红外小目标在第k帧图像中的真实位置进行级联匹配，得到初次匹配成功的结果、初次未匹配的轨迹以及初次未匹配的检测框。具体地，计算基于余弦距离的代价矩阵，以删除余弦矩阵过大的检测框，并使用匈牙利算法将卡尔曼预测轨迹和实际检测框之间的马氏距离为代价矩阵，将轨迹和检测框进行匹配，返回匹配结果。需要说明的是，由于匈牙利算法本身具有延时性，并且为了克服干扰，步骤S42的级联匹配在循环检测30帧之间进行，若30帧后仍然未匹配成功，则放弃匹配。

[0089] 步骤S43，对步骤S42中初次未匹配的轨迹和初次未匹配的检测框进行 IOU匹配，得到再次匹配成功的结果、再次未匹配的轨迹以及再次未匹配的检测框。具体地，对只有一帧匹配的轨迹，计算其与未匹配检测框之间的IOU 距离，并删除IOU距离过大的检测框，并使用匈牙利算法将卡尔曼预测轨迹和实际检测框之间的IOU距离为代价矩阵，将轨迹和检测框进行匹配，返回匹配结果。需要说明的是，步骤S43的IOU匹配在每一帧图像之间进行，即，若在当前时刻因为干扰等原因未匹配成功，再下一帧不会再对当前帧的内容进行匹配。

[0090] 步骤S44，根据初次匹配成功的结果和再次匹配成功的结果，更新卡尔曼滤波的参数。具体地，采用上述卡尔曼滤波公式对卡尔曼滤波参数中的协方差和均值进行更新，P即为协方差，$\hat{x}$是轨迹的预测值，其期望$E(\hat{x})$即为均值。

[0091] 步骤S45，为再次未匹配的检测框分配新的轨迹和新的ID，并通过ReID提取检测框中目标物的特征集；同时判断再次未匹配的轨迹是否处于确定状态，保留处于确定状态且失配数小于30帧的轨迹，重复步骤S41-步骤S44。具体来说，若检测框没有和现存的跟踪轨迹匹配上，则对该检测框进行跟踪轨迹初始化(采用是余弦相似度量矩阵进行初始化，默认值为0)。新的轨迹初始化的状态是未确定态，只有满足连续三帧都成功匹配，才能将未确定状态转化为确定状态。若确定状态的轨迹必须和目标物连续失配超过30次，才会被删除。

[0092] Deepsort模型利用匈牙利算法匹配运动的目标，并使用深度识别网络识别目标标

识，依据目标标识追踪目标，同时将目标标识信息存储，当目标被丢失一段时间又重现后，可以根据目标标识重新识别出。匈牙利算法是用以解决检测结果与追踪预测结果的关联的方法，使用对已存在的运动目标的运动状态的卡尔曼预测结果与检测结果之间的马氏距离进行运行信息的关联。

[0093] 马氏距离公式如下：

[0094] $d^{(1)}(i,j)=(d_i-y_i)^T S_i^{-1}(d_i-y_i)$

[0095] d_i表示第i个检测位置，y_i表示第i个追踪器对目标的预测位置，S_i表示检测位置与平均追踪位置之间的协方差矩阵；

[0096] 马氏距离通过计算检测位置和平均追踪位置之间的标准差将状态测量的不确定性进行了考虑。如果某次关联的马氏距离小于指定的阈值$t^{(1)}$，则设置运动状态关联成功，使用的函数为：

[0097] $$b_{i,j}^{(1)}=\mathbb{1}[d^{(1)}(i,j)\le t^{(1)}]$$

[0098] $b_{i,j}^{(1)}$为1时表示关联成功，$t^{(1)}$取9.4877。

[0099] 步骤S5，利用红外小目标检测与跟踪的识别器，对红外小目标进行实时检测跟踪。

[0100] 具体地，步骤S5包括：

[0101] 步骤S51，设定红外小目标检测与跟踪的识别器中的参数，并将 YOLOv5s模型的权重参数加载至YOLOv5s模型，将Deepsort模型的权重参数加载至Deepsort模型，得到YOLOv5s模型和Deepsort模型的串行模型。其中，YOLOv5s模型的置信度阈值设为0.01，交并比设为0.4。

[0102] 步骤S52，利用红外成像设备采集红外小目标的实时图像，并对采集到的实时图像进行对比度增强处理，获取增强后的实时图像，以降低红外小目标的特征识别难度。

[0103] 对采集到的实时图像进行对比度增强处理的方法为：获取每一帧图像的直方图分布，将每帧图像的直方图分布改变成近似均匀分布直方图，以增强图像的对比度。

[0104] 将直方图分布改变成近似均匀分布直方图的方法为：对原图像的每个像素点进行单调非线性映射变换f，即$D_B=f(D_A)$，并保持像素点总数不变，如下：$\int_0^{D_A}H_A(D)\,dD=\int_0^{D_B}H_B(D)\,dD$，其中$H_A(D)$为原图像直方分布，$H_B(D)$为均匀分布，取$H_B(D)=\frac{A_0}{L}$，$A_0$为像素点个数，L为灰度级深度，取256。

[0105] 步骤S53，对增强后的实时图像的前100帧进行像素平均统计，获取实时图像中的盲点，将每个盲点的位置存入盲点列表。具体地，盲点的获取方法为：将前100帧图像在同一个点的像素灰度进行平均，由于盲点像素的平均值远高于正常点的平均值，截取盲点像素平均值的0.7倍～0.9倍作为阈值，超过阈值的像素视为盲点。

[0106] 步骤S54，将增强后的实时图像输入YOLOV5s模型，生成置信度在0.01以上的实时目标候选框集。

[0107] 步骤S55，将实时目标候选框集输入至Deepsort模型，获取滤波后的实时目标候选框集，并剔除滤波后的实时目标候选框集中包含盲点列表的目标候选框，生成最终目标检测框，实现红外小目标的实时跟踪。

[0108] 对红外小目标进行实时检测跟踪，可在验证集上计算精确率、召回率、F1分数、mAP

等。图4(a)-图4(h)是置信度阈值为0.001下的上述各性能指标图，从图中可以看出，在给定阈值0.001下，大致能达到精确率为 0.94，召回率为0.75，图4(a)表示标注框和预测框产生的误差。出现跳变的原因是在训练途中中断，并调整了标签。其中召回率较低，是因为图片所给的训练集部分情况被目标背景所遮挡或被背景光所淹没，这些情况需要用预测类的算法和对目标背景和目标的分离算法来进行处理。

[0109] 图5(a)是本发明识别检测模型测试集上的P-R曲线图。P-R曲线能够直观的反映模型的性能，在P-R曲线中，以精确率为纵坐标，召回率为横坐标，并且PR曲线下面的面积越大，模型的性能越好。从图中来看，模型的性能较好。

[0110] 图5(b)是不同置信度阈值下，F1分数的变化图。由于精确率和召回率都反映模型的好坏，但是需要同时对这两个量进行衡量，采用F1分数曲线，其中F1分数是精确率和召回率的调和平均数，在置信度阈值较小时，随着置信度阈值的增加，F1性能指标增加；随后在置信度阈值增加时，F1性能指标下降。

[0111] 图5(c)是不同置信度阈值下，对应的精确率变化图。当置信度为 0.1时，预测精确度基本收敛到99％。

[0112] 图5(d)是不同置信度阈值下，对应的召回率变化图。随着置信度升高，召回率不断下降。

[0113] 最终跟踪预测效果如图6(a)-图6(c)所示，分别存在单目标和多目标的情况，锚框上的数字是目标物的id号，用于跟踪记录。

[0114] 另外，置信度阈值变化如表1所示：

表1置信度变化对比表

置信度阈值	精确率	召回率	mAP@0.5	mAP@[0.5:0.95]
0.001	0.937	0.756	0.786	0.635
0.01	0.937	0.756	0.771	0.625
0.1	0.998	0.5	0.505	0.434

[0115] 本发明结合了用于目标检测的YOLOv5s模型和用于目标跟踪的 Deepsort模型，并针对红外小目标场景下的复杂背景设计了数据清洗及增强等预处理和背景去除的后处理，能够准确且快速地检测并跟踪复杂背景下的红外小目标。

[0116] 以上所述的，仅为本发明的较佳实施例，并非用以限定本发明的范围，本发明的上述实施例还可以做出各种变化。即凡是依据本发明申请的权利要求书及说明书内容所作的简单、等效变化与修饰，皆落入本发明专利的权利要求保护范围。本发明未详尽描述的均为常规技术内容。

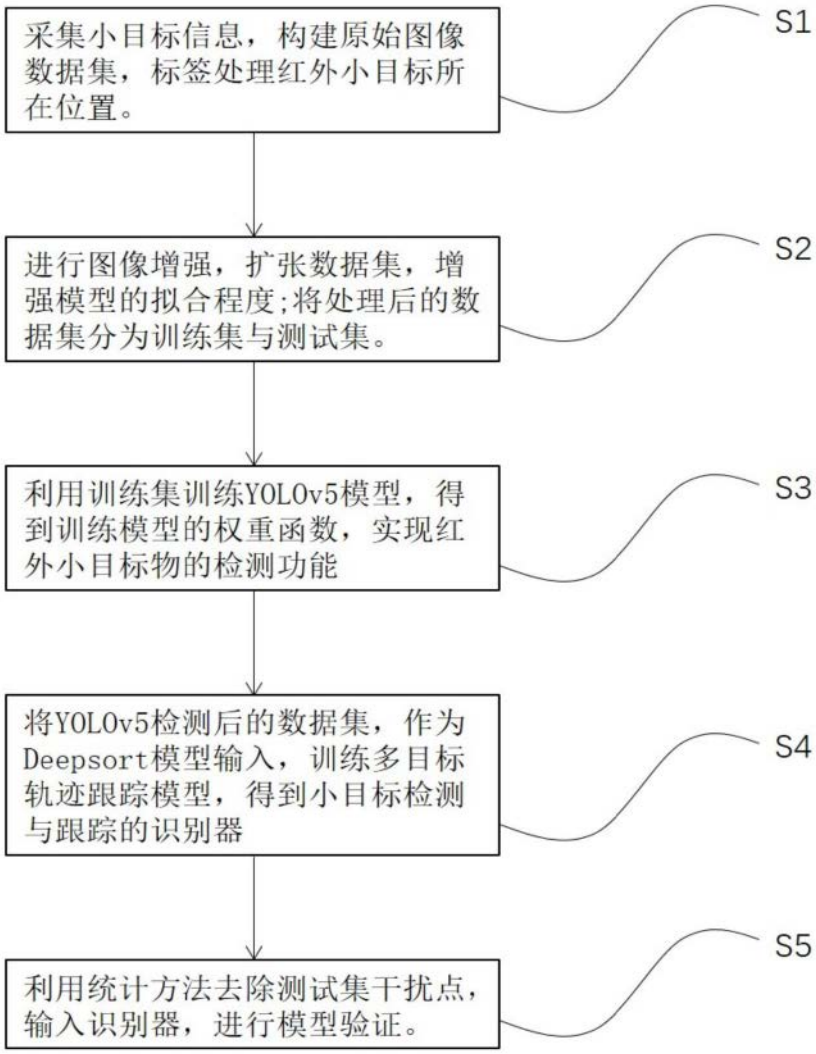

采集小目标信息，构建原始图像数据集，标签处理红外小目标所在位置。
S1
进行图像增强，扩张数据集，增强模型的拟合程度;将处理后的数据集分为训练集与测试集。
S2
利用训练集训练YOLOv5模型，得到训练模型的权重函数，实现红外小目标物的检测功能
S3
将YOLOv5检测后的数据集，作为Deepsort模型输入，训练多目标轨迹跟踪模型，得到小目标检测与跟踪的识别器
S4
利用统计方法去除测试集干扰点，输入识别器，进行模型验证。
S5

图1

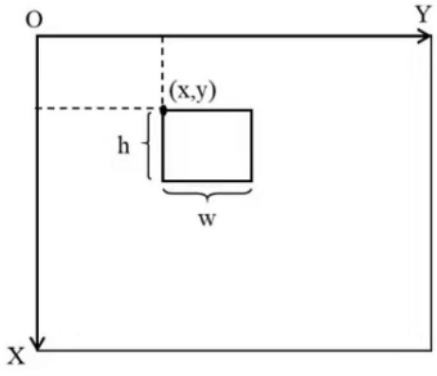

O
Y
(x,y)
h
w
X

图2

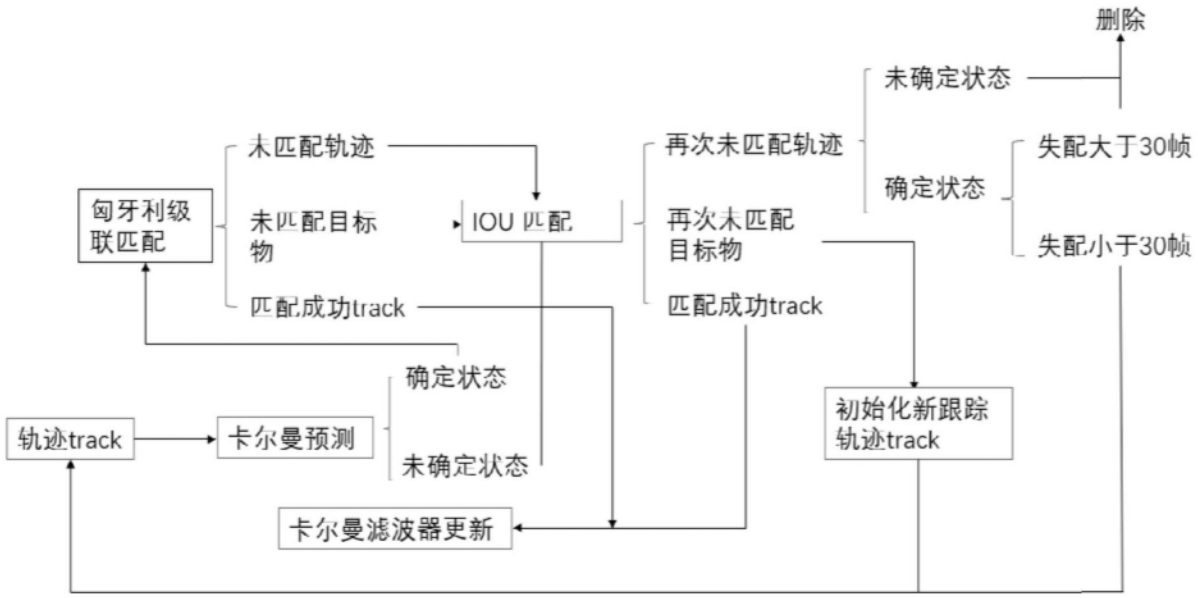

删除
未确定状态
失配大于30帧
确定状态
失配小于30帧
未匹配轨迹
匈牙利级联匹配
未匹配目标物
IOU 匹配
再次未匹配轨迹
再次未匹配目标物
匹配成功track
匹配成功track
确定状态
初始化新跟踪轨迹track
轨迹track
卡尔曼预测
未确定状态
卡尔曼滤波器更新

图3

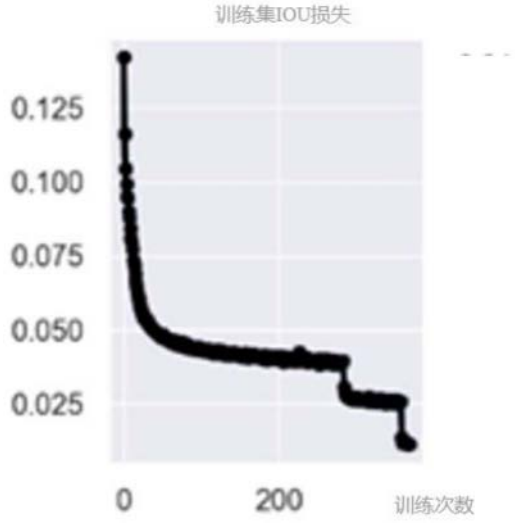

训练集IOU损失
0.125
0.100
0.075
0.050
0.025
0
200
训练次数

图4(a)

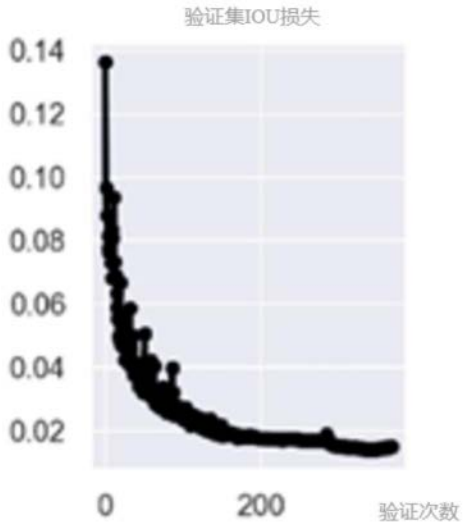

验证集IOU损失
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
200
验证次数

图4(b)

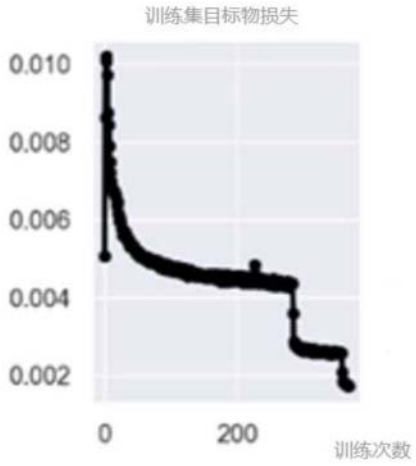

训练集目标物损失
0.010
0.008
0.006
0.004
0.002
0
200
训练次数

图4(c)

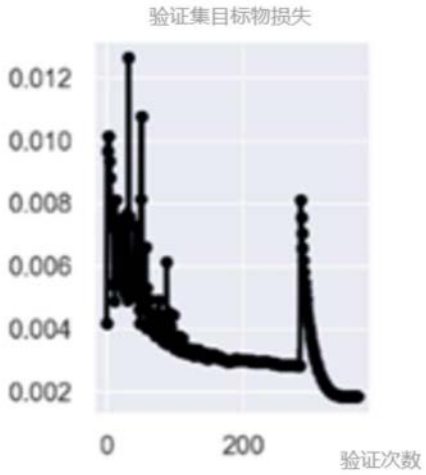

验证集目标物损失
0.012
0.010
0.008
0.006
0.004
0.002
0
200
验证次数

图4(d)

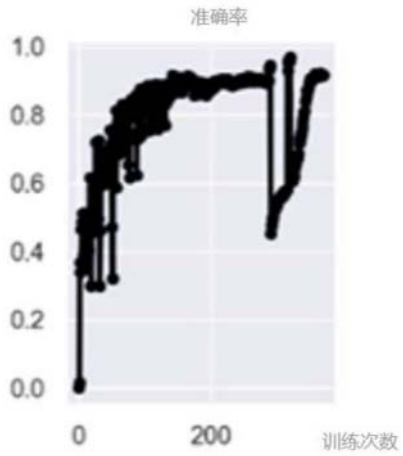

准确率
1.0
0.8
0.6
0.4
0.2
0.0
0
200
训练次数

图4(e)

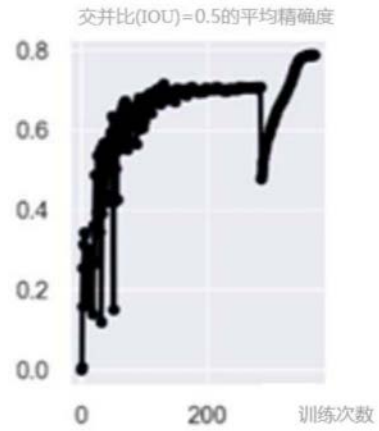

交并比(IOU)=0.5的平均精确度
0.8
0.6
0.4
0.2
0.0
0
200
训练次数

图4(f)

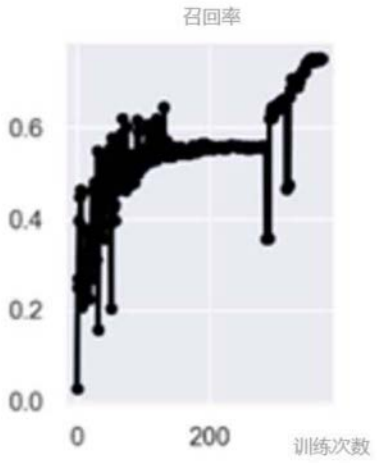

召回率
0.6
0.4
0.2
0.0
0
200
训练次数

图4(g)

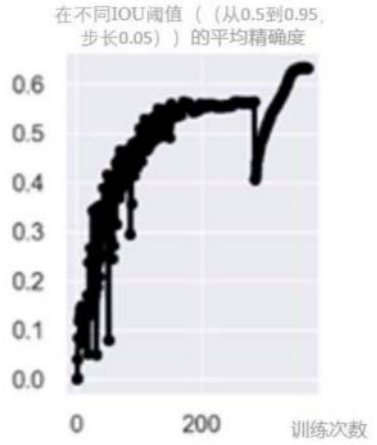

在不同IOU阈值（（从0.5到0.95，步长0.05））的平均精确度
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
200
训练次数

图4(h)

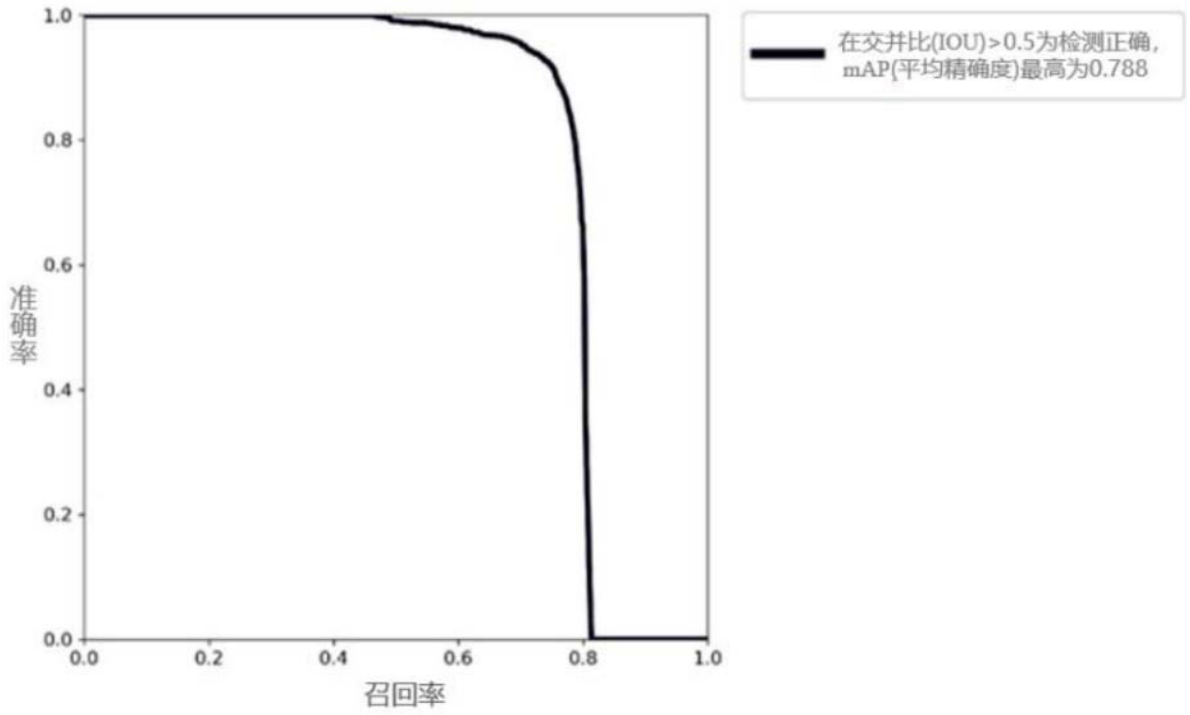

在交并比(IOU)>0.5为检测正确，
mAP(平均精确度)最高为0.788
准确率
召回率
0.0
0.2
0.4
0.6
0.8
1.0

图5(a)

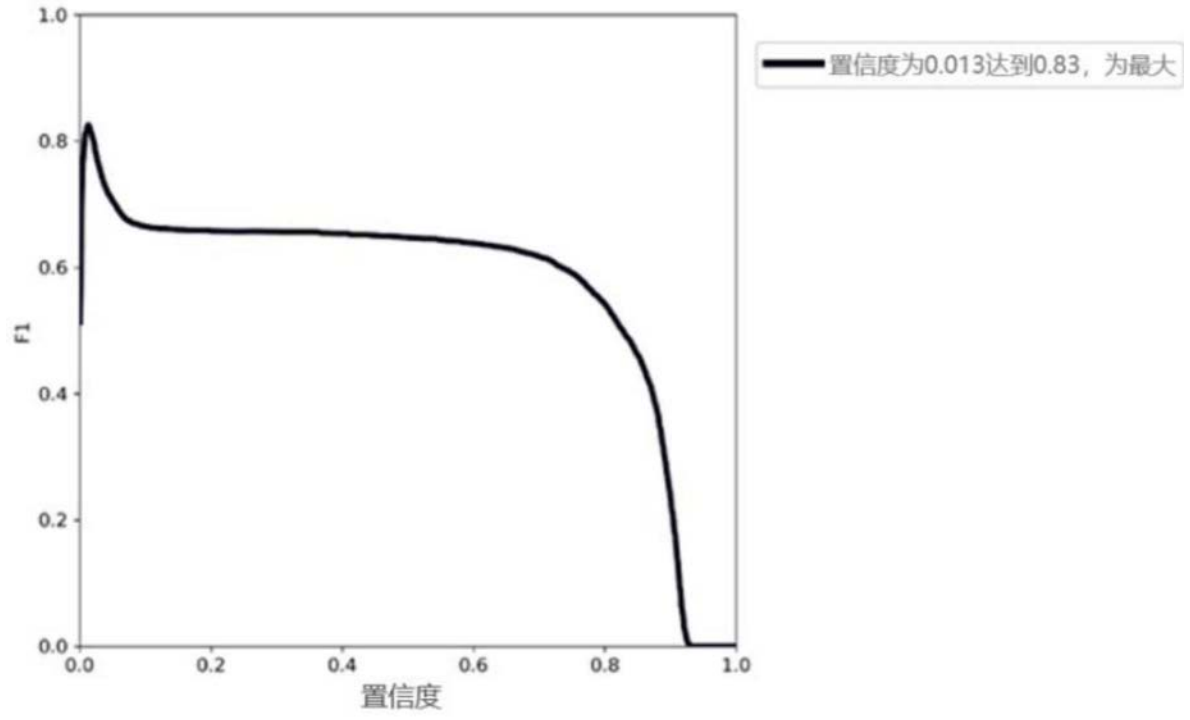

置信度为0.013达到0.83，为最大
F1
置信度
0.0
0.2
0.4
0.6
0.8
1.0

图5(b)

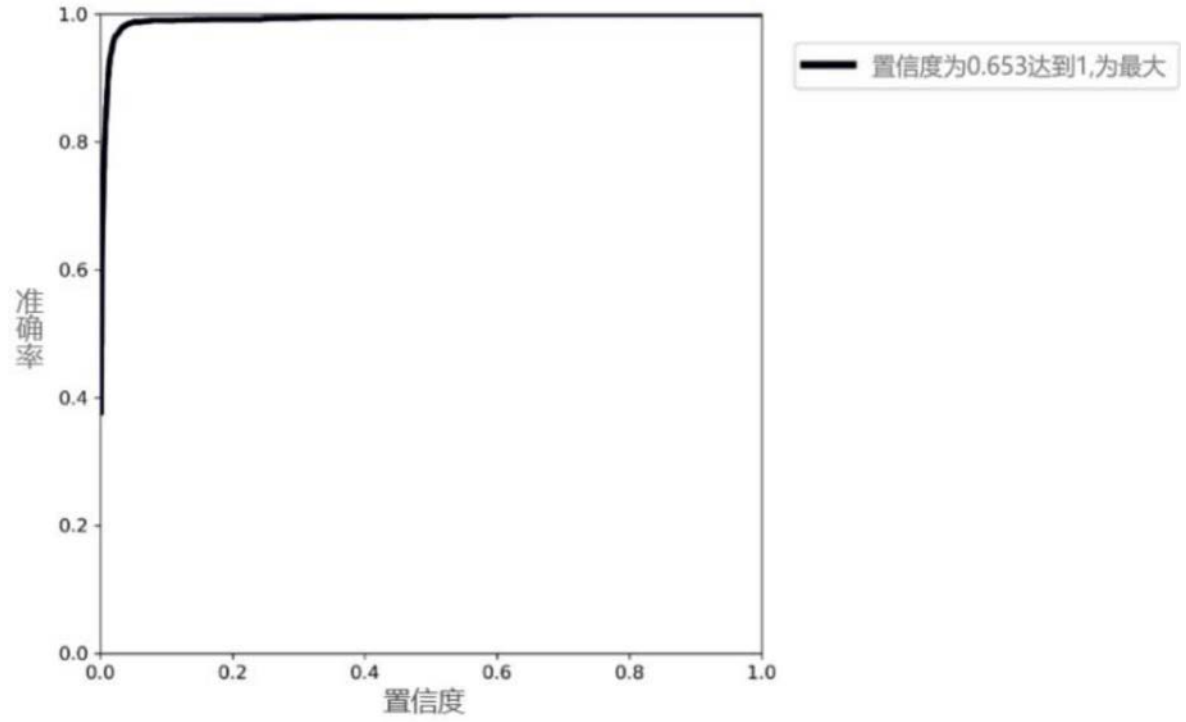

置信度为0.653达到1,为最大
准确率
置信度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

图5(c)

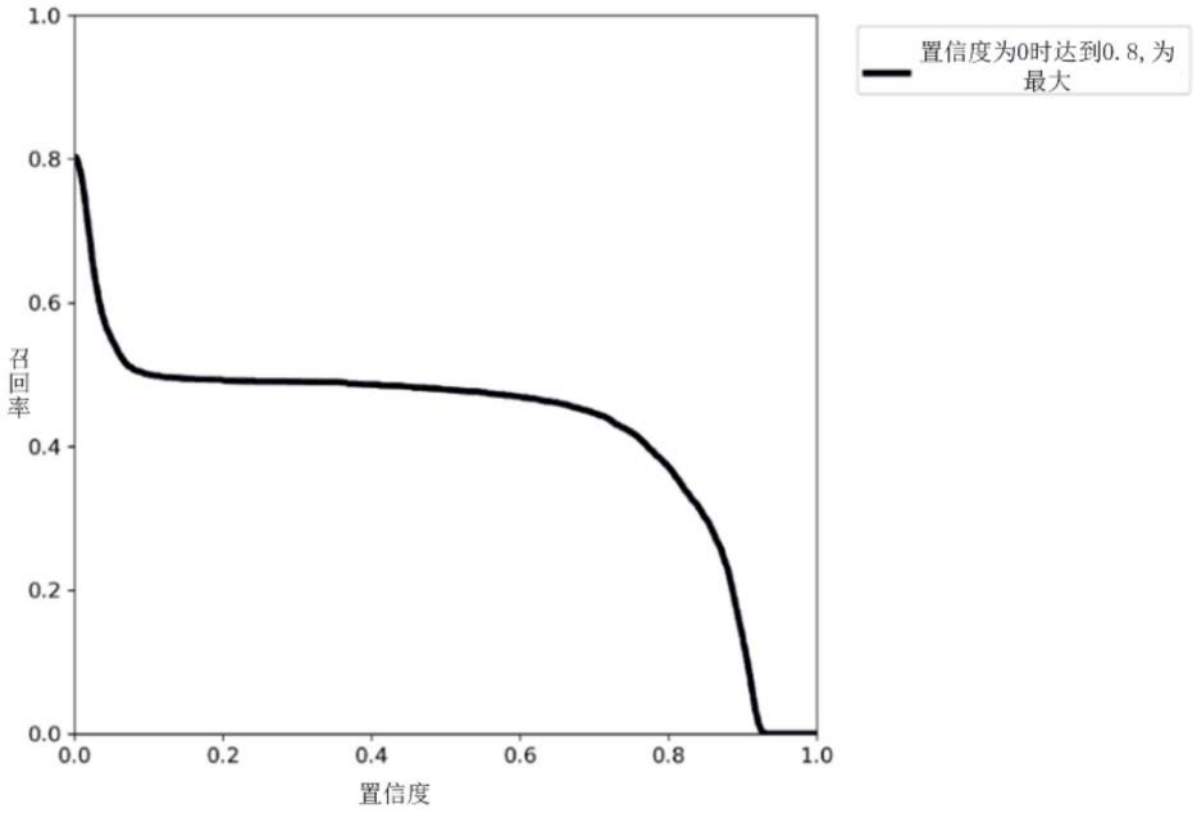

置信度为0时达到0.8,为最大
召回率
置信度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

图5(d)

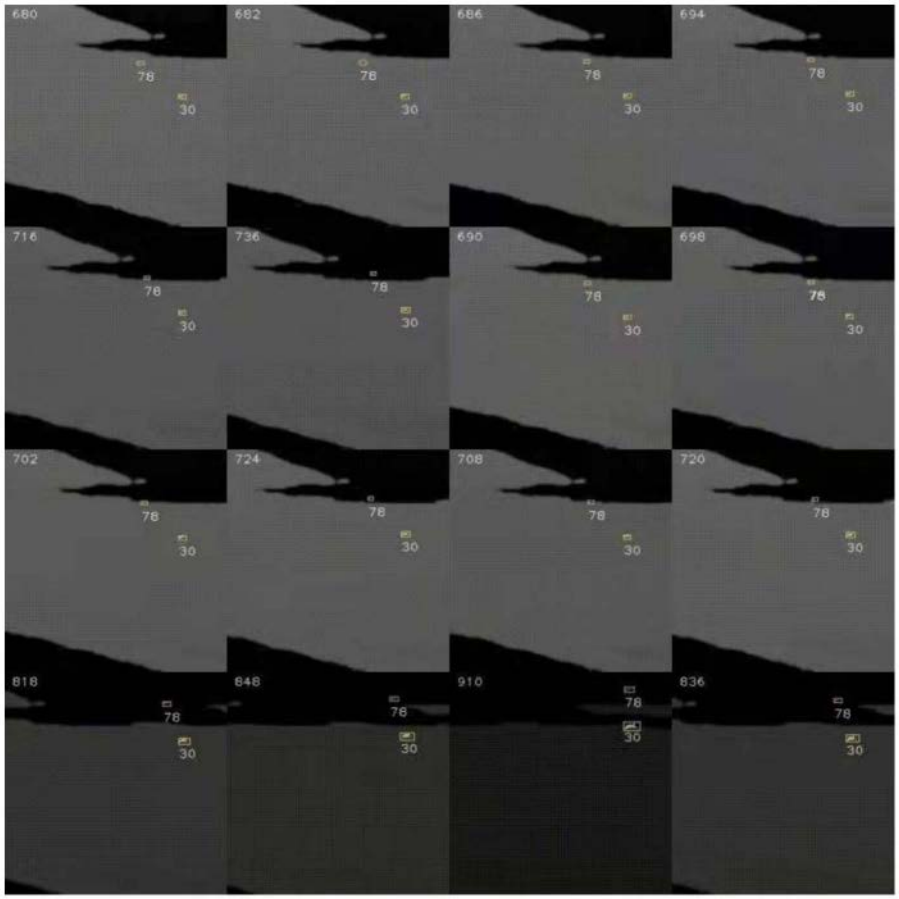

680
682
686
694
716
736
690
698
702
724
708
720
818
848
910
836
78
30

图6(a)

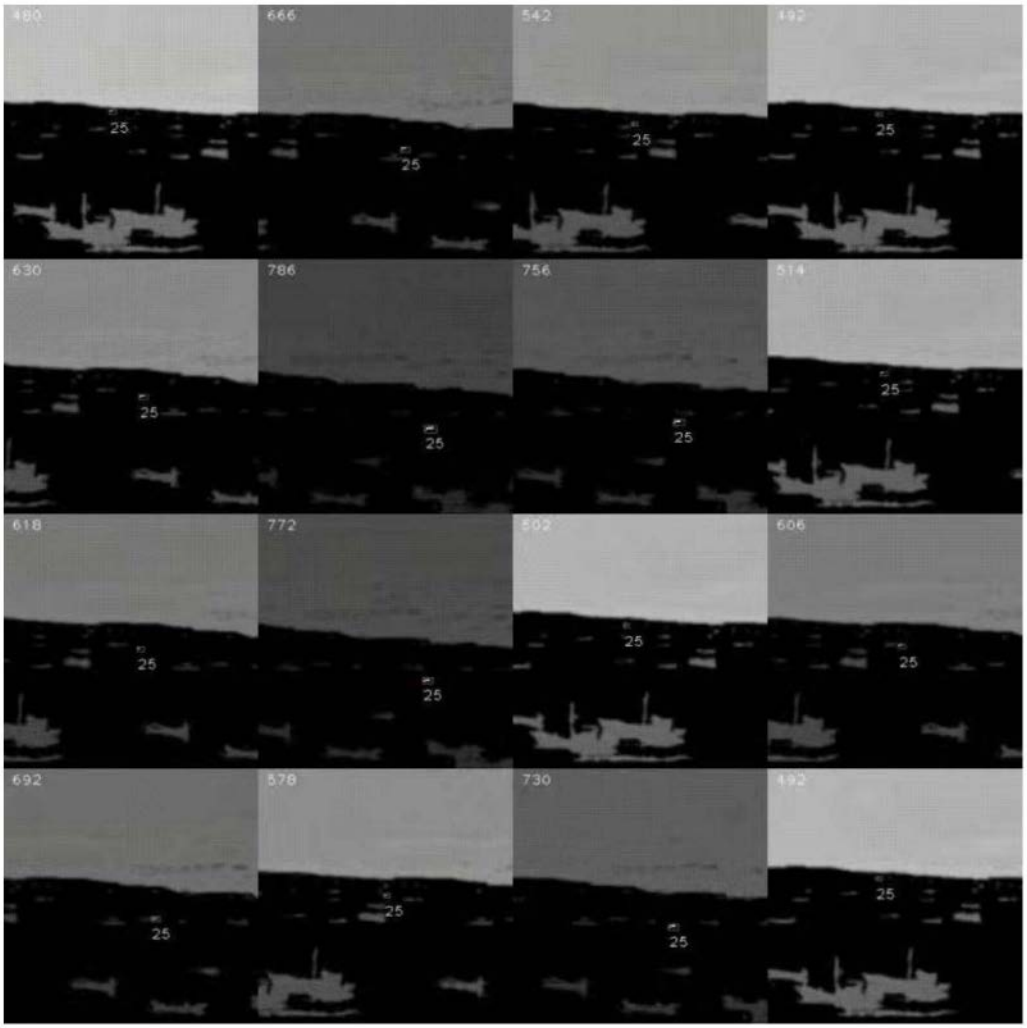


图6(b)

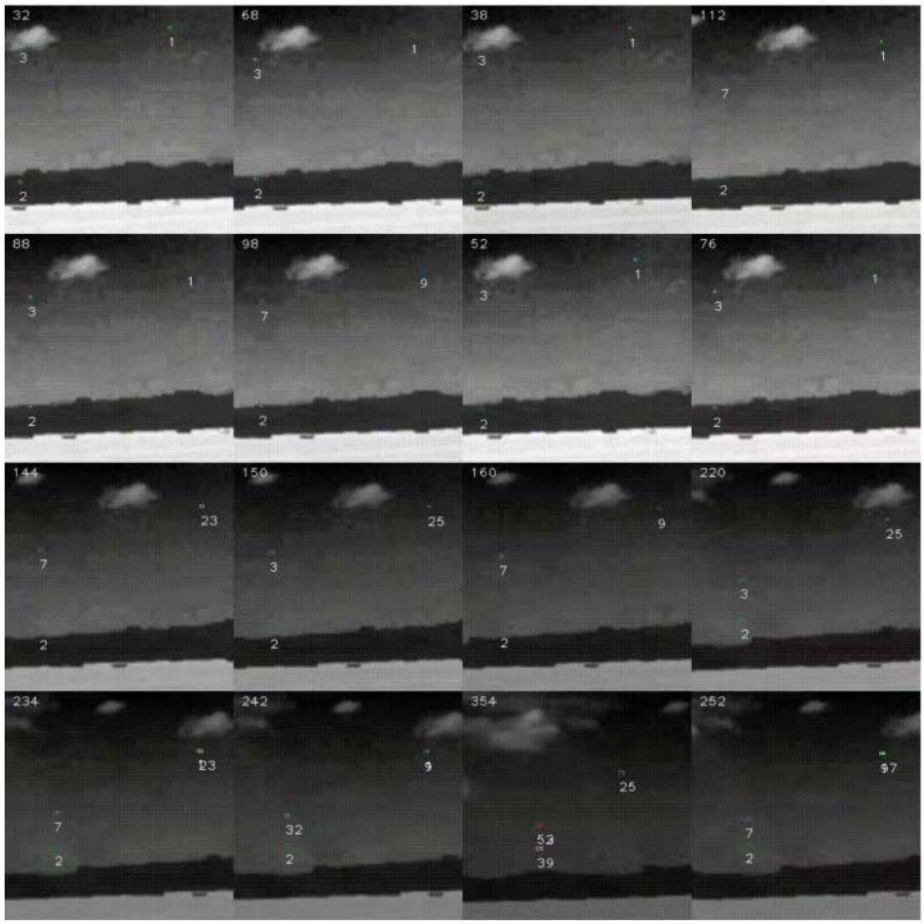


图6(c)