



(12) 发明专利申请

(10) 申请公布号 CN 114529799 A

(43) 申请公布日 2022.05.24

(21) 申请号 202210009200.8

G06N 3/04 (2006.01)

(22) 申请日 2022.01.06

G06N 3/08 (2006.01)

(71) 申请人 浙江工业大学

G06T 7/246 (2017.01)

地址 310014 浙江省杭州市拱墅区潮王路
18号

G06T 7/277 (2017.01)

(72) 发明人 周德龙 任威

(74) 专利代理机构 杭州天正专利事务所有限公
司 33201

专利代理师 王兵 舒良

(51) Int. Cl.

G06V 20/00 (2022.01)

G06V 10/44 (2022.01)

G06V 10/764 (2022.01)

G06V 10/82 (2022.01)

G06K 9/62 (2022.01)

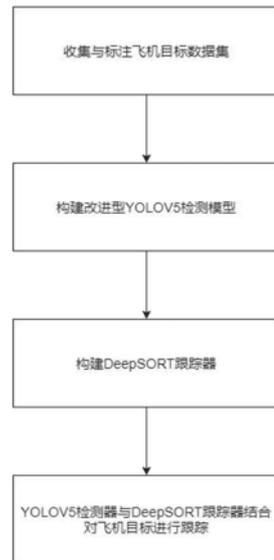
权利要求书3页 说明书7页 附图4页

(54) 发明名称

一种基于改进型YOLOV5算法的飞机多目标跟踪方法

(57) 摘要

一种基于改进型YOLOV5算法的飞机多目标跟踪方法,具体步骤如下:步骤1:收集与标注飞机目标数据集;步骤2:基于步骤1中的数据集对YOLOV5目标检测网络进行训练,并针对飞机目标进行改进;步骤3:构建DeepSORT跟踪器;步骤4:基于YOLOV5检测器与DeepSORT跟踪器结合进行飞机目标跟踪。本发明针对飞机多目标跟踪中存在的小目标漏检的问题,对YOLOV5检测模型进行改进,能够实现对多场景下飞机目标的检测与多目标跟踪。



1. 一种基于改进型YOLOV5算法的飞机多目标跟踪方法,其特征在于,包括以下步骤:

步骤1:收集与标注飞机目标数据集,具体包括:

收集飞机图像数据以及视频数据,使用标注工具Labelimg、Darklabel分别对图像和视频进行标注,并构建飞机图像数据集,将所述数据集划分成训练集、测试集以及验证集;

步骤2:基于步骤1中的数据集对YOLOV5目标检测网络进行训练,并针对飞机目标进行改进,具体包括:

步骤2.1:基于BiFPN结构进行多特征融合:

YOLOV5的主干网络由CONV模块、C3模块、SPPF模块结合构成,可以产生多尺度的特征图,利用BiFPN结构在head部分中加入多特征融合,针对飞机小目标的漏检问题,使用4种不同尺度下的特征检测图层,对大小不同的飞机目标进行检测;

步骤2.2:使用CBAM注意力机制模块增加YOLOV5模型检测精准度:

使用CBAM注意力机制,将通道注意力机制与空间注意力机制融合,使特征能够覆盖到检测物体的更多部位,最终使判断物体的几率提高,增强模型的检测精准度;

步骤2.3:获得YOLOV5检测模型:

YOLOV5网络包括主干网络backbone部分、head部分,首先对训练集中的图像数据预处理将所有图像统一分辨率,并将预处理后的图像输入到backbone部分中,获得不同尺度的特征图,然后将不同尺寸的特征图输入到head部分,在head部分中对特征图依次进行上采样、下采样以及基于通道维度的拼接操作产生多种尺度的特征图,并利用步骤2.1中的特征融合操作与步骤2.2的CBAM注意力机制获得不同尺度的特征图并输入到head部分中的Detect层,再将多种损失函数结合并进行反向传播对模型中的梯度进行更新调节权重参数,最后使用验证集对现有模型进行验证,获得YOLOV5目标检测模型;

步骤3:构建DeepSORT跟踪器,具体包括:

步骤3.1:构建DeepSORT跟踪器运动模型:

使用卡尔曼滤波对已检测到的飞机目标构建运动模型并进行预测,预测下一帧中飞机目标位置;

步骤3.2:构建DeepSORT跟踪器特征提取模块:

使用步骤1中Darklabel对视频进行标注的图像数据,并对图像进行裁剪,对所得到的飞机图像数据进行分类,最后对数据集划分训练集和验证集;将所得数据输入到ResNeXt50网络中,并得到ResNeXt50特征提取模型,并使用该特征检测模型对YOLOV5所检测到的飞机目标进行特征提取;

步骤3.3:针对DeepSORT跟踪器中目标前后帧跟踪轨迹与检测目标之间规定距离度量标准:

首先使用马氏距离表示运动模型之间的关联度并作为距离函数:

$$d^{(1)}(i, j) = (d_j - y_j)^T S_i^{-1} (d_j - y_j) \quad (1)$$

其中 d_j 表示第j个YOLO V5检测到的目标位置, y_i 表示第i个DeepSORT跟踪器中所预测的目标位置, S_i^{-1} 表示跟踪到的目标位置与检测目标位置的协方差矩阵的逆矩阵, $d^{(1)}$ 表示在多维运动状态数据中YOLOV5所检测到目标的检测框与卡尔曼滤波所得到的预测框的关联程度;

再加入特征之间的余弦距离,对检测到的目标特征与跟踪到的目标特征之间计算余弦距离:

$$d^{(2)}(i, j) = \min\{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in \mathfrak{R}_i\} \quad (2)$$

其中 r_j 表示第 j 个检测到所提取的目标特征并添加限制条件 $\|r_j\|=1$ 所得, \mathfrak{R}_i 为每个跟踪目标构建一个轨迹,在其中存储每个跟踪目标在最近的 M 帧中关联成功的特征向量, $r_k^{(i)}$ 表示第 k 帧中第 i 个跟踪到的目标特征, $d^{(2)}$ 表示第 i 个跟踪到的最近 N 个成功关联的特征集与当前帧中YOLOV5检测到的目标特征之间的最小余弦距离,并对以上两种方法设定阈值,设定匹配时的度量标准,最终度量方式由二者的加权求和所得,公式如下:

$$c_{i,j} = \lambda d^{(1)}(i, j) + (1 - \lambda) d^{(2)}(i, j) \quad (3)$$

步骤3.4:基于匈牙利算法对多个检测目标与跟踪轨迹进行匹配:

使用匈牙利算法对DeepSORT跟踪器中的跟踪轨迹与YOLOV5当前所检测到的目标进行匹配,利用步骤3.2中的特征提取模块,以及步骤3.3中的距离度量标准组合成级联匹配,对跟踪轨迹与检测目标之间进行距离计算,并设定阈值过滤匹配结果,之后还需要在IOU匹配中使用匈牙利算法对跟踪轨迹与检测目标再进行一次匹配,最后利用匹配到的结果对当前所有轨迹分成三个状态确认态、未确认态、删除态;

步骤4:基于YOLOV5检测器与DeepSORT跟踪器结合进行飞机目标跟踪,具体包括:

从摄像头或者视频流中读入第一帧,第一帧中使用YOLOV5检测器对帧中的飞机物体进行识别,将所检测到的飞机目标的坐标信息、置信度以及飞机目标图像信息传递给DeepSORT跟踪器中,并设置置信度阈值对检测框进行过滤,删除置信度过小的检测框,第一帧中无跟踪目标,利用YOLOV5检测器对目标进行检测并在DeepSORT跟踪器中创建新的Tracker,将检测到的目标存入tracks中,并初始化卡尔曼滤波器状态,Tracker中的tracks状态初始化为未确认态;

读入第二帧,YOLOV5检测后将检测信息输入到DeepSORT跟踪器中,利用第一帧中创建的Tracker中的卡尔曼滤波器以及YOLOV5的坐标信息,获得当前帧中的预测信息,由于第一帧中无确认态的tracks,则对未匹配的tracks以及未匹配的检测目标再进行IOU匹配,计算出IOU最大的唯一匹配,最后对Tracker中各个tracks的状态进行更新,对于已匹配的tracks将所匹配目的特征保存到各个tracks的特征集中并更新卡尔曼滤波状态,对于未匹配的检测目标创建新的tracks并进行初始化,对于依然未匹配到的tracks直接删除;

读入第三帧,YOLOV5检测后将检测信息输入到DeepSORT跟踪器中,该帧中依然没有确认态的tracks,在完成前后两帧检测目标与跟踪轨迹之间的IOU匹配后对已经连续三帧确认匹配的tracks变更为确定态,对确定态的tracks进行特征更新,并对卡尔曼滤波状态进行更新,对于依然未匹配到的tracks直接删除,对未匹配到的检测目标创建新的tracks并进行初始化;

读入第四帧,YOLOV5检测后将检测信息输入到DeepSORT跟踪器中,并使用步骤3.2中的特征提取模块对YOLOV5检测到的目标进行特征提取,并利用tracks中的卡尔曼滤波信息以及特征信息,以及步骤3.4中马氏距离、余弦距离以及匈牙利算法组合的级联匹配算法,将处于确认态的tracks与当前帧中所提取到的目标特征进行匹配,对已匹配的之前为确认态的tracks更新特征集以及卡尔曼滤波状态,对已匹配的之前为未确认态的tracks匹配命中

次数增加1,对未匹配的tracks以及未匹配的检测目标再进行IOU匹配,如果tracks仍未匹配且之前为确认态,需要经过max_age帧后依然没有完成匹配,则将tracks删除,如果tracks仍未匹配且之前的状态为未确认态,则直接删除tracks,对未匹配的到检测目标创建新的tracks并进行初始化,在之后的每一帧中都与第四帧相同的方法进行匹配,最后将跟踪框结果输出到视频中,并将视频中的所跟踪到的飞机目标的坐标信息输出。

2.如权利要求1所述的一种基于改进型YOLOV5算法的飞机多目标跟踪方法,其特征在于,步骤3.3中的M和N都取值为70,步骤4的置信度阈值取值为0.3。

一种基于改进型YOLOV5算法的飞机多目标跟踪方法

技术领域

[0001] 本发明可以实现对多种类飞机的多目标的实时检测与跟踪,其中针对多目标跟踪中飞机的小型目标容易漏检的问题提出了解决方法,针对YOLOV5的网络结构进行了针对性的改进,改善飞机在多目标跟踪时的检测精度进而改善整个跟踪方法的精准度。DeepSORT跟踪器与YOLOV5检测器的结合能够通过前后帧检测与预测的结合实现对多目标的实时跟踪与检测,通过改善YOLO V5检测器的性能能够增强跟踪效果。YOLOV5检测器可以使用较小的模型能够快速且精准的对目标进行检测。本发明能够对机场的民航飞机、航展中各式飞机进行实时检测和跟踪,能够对辅助完成机场对飞机的精准调度以及能够辅助战斗机对各式飞机进行跟踪。

背景技术

[0002] 实时多目标跟踪是在计算机视觉领域的一个重要分支,有很广泛的应用范围与价值。随着硬件设备的逐步更新,伴随着人工智能的兴起,出现越来越多的基于深度学习的跟踪算法。

[0003] 目标跟踪对视频中的图像按照时间序列对感兴趣的运动目标进行检测,并提取器表观特征,以及进行匹配最后完成跟踪。多目标跟踪主要任务是对视频中的感兴趣的的目标确定其位置,给其分配ID并一直维持以及记录目标的运动轨迹。

[0004] 首先,针对机场的特殊环境会出现很多遮挡现象,导致飞机目标难以识别,以及摄像头的远近也影响着检测性能,出现过小飞机目标时会出现无法检测的情况;其次传统的目标检测算法在该情况下无法精准检测小型飞机目标,传统目标检测算法难以检测出现部分遮挡和环境光变化下的目标;最后飞机的多目标跟踪主要取决于检测器对目标检测的准确性,目标检测的精度会直接影响多目标跟踪的效果。在机场环境下利用摄像头读入图像与飞机多目标跟踪方法结合,可以对每架飞机进行精准检测与跟踪,而且能够保证跟踪的实时性,能够协助完成对每架飞机的调度。

发明内容

[0005] 本发明要克服现有技术的上述缺点,提供一种基于改进型YOLO V5算法的飞机多目标检测方法,改进YOLOV5网络结构并使用在飞机多目标跟踪中。

[0006] 为了实现以上目的,本发明所采用的技术方案是:

[0007] 一种基于改进型YOLOV5算法的飞机多目标跟踪方法,其特征在于,包括以下步骤:

[0008] 步骤1:收集与标注飞机目标数据集,具体包括:

[0009] 收集飞机图像数据以及视频数据,使用标注工具Labelimg、Darklabel分别对图像和视频进行标注,并构建飞机图像数据集,将所述数据集划分成训练集、测试集以及验证集;

[0010] 步骤2:基于步骤1中的数据集对YOLOV5目标检测网络进行训练,并针对飞机目标进行改进,具体包括:

[0011] 步骤2.1:基于BiFPN结构进行多特征融合:

[0012] YOLOV5的主干网络由CONV模块、C3模块、SPPF模块结合构成,可以产生多尺度的特征图,利用BiFPN结构在head部分中加入多特征融合,针对飞机小目标的漏检问题,使用4种不同尺度下的特征检测图层,对大小不同的飞机目标进行检测。

[0013] 步骤2.2:使用CBAM注意力机制模块增加YOLOV5模型检测精准度:

[0014] 使用CBAM注意力机制,将通道注意力机制与空间注意力机制融合,使特征能够覆盖到检测物体的更多部位,最终使判断物体的几率提高,增强模型的检测精准度。

[0015] 步骤2.3:获得YOLOV5检测模型:

[0016] YOLOV5网络包括主干网络backbone部分、head部分,首先对训练集中的图像数据预处理将所有图像统一分辨率,并将预处理后的图像输入到backbone部分中,获得不同尺度的特征图,然后将不同尺寸的特征图输入到head部分,在head部分中对特征图依次进行上采样、下采样以及基于通道维度的拼接操作产生多种尺度的特征图,并利用步骤2.1中的特征融合操作与步骤2.2的CBAM注意力机制获得不同尺度的特征图并输入到head部分中的Detect层,再将多种损失函数结合并进行反向传播对模型中的梯度进行更新调节权重参数,最后使用验证集对现有模型进行验证,获得YOLOV5目标检测模型。

[0017] 步骤3:构建DeepSORT跟踪器对已检测的目标进行预测与跟踪,具体包括:

[0018] 步骤3.1:构建DeepSORT跟踪器运动模型:

[0019] 使用卡尔曼滤波对已检测到的飞机目标构建运动模型并进行预测,预测下一帧中飞机目标位置。

[0020] 步骤3.2:获得DeepSORT跟踪器特征提取模块:

[0021] 使用步骤1中Darklabel对视频进行标注的图像数据,并对图像进行裁剪,对所得到的飞机图像数据进行分类,最后对数据集划分训练集和验证集。将所得数据输入到ResNeXt50网络中,并得到ResNeXt50特征提取模型,并使用该特征检测模型对YOLOV5所检测到的飞机目标进行特征提取。

[0022] 步骤3.3:针对DeepSORT跟踪器中目标前后帧跟踪轨迹与检测目标之间规定距离度量标准:

[0023] 首先使用马氏距离表示运动模型之间的关联度并作为距离函数:

$$[0024] \quad d^{(1)}(i, j) = (d_j - y_j)^T S_i^{-1} (d_j - y_j) \quad (1)$$

[0025] 其中 d_j 表示第j个YOLO V5检测到的目标位置, y_i 表示第i个DeepSORT跟踪器中所预测的目标位置, S_i^{-1} 表示跟踪到的目标位置与检测目标位置的协方差矩阵的逆矩阵, $d^{(1)}$ 表示在多维运动状态数据中YOLOV5所检测到目标的检测框与卡尔曼滤波所得到的预测框的关联程度。

[0026] 再加入特征之间的余弦距离,对检测到的目标特征与跟踪到的目标特征之间计算余弦距离:

$$[0027] \quad d^{(2)}(i, j) = \min \{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in \mathfrak{R}_i\} \quad (2)$$

[0028] 其中 r_j 表示第j个检测到所提取的目标特征并添加限制条件 $\|r_j\| = 1$ 所得, \mathfrak{R}_i 为每个跟踪目标构建一个轨迹,在其中存储每个跟踪目标在最近的M帧中关联成功的特征向

量, $r_k^{(i)}$ 表示第k帧中第i个跟踪到的目标特征, $d^{(2)}$ 表示第i个跟踪到的最近N个成功关联的特征集与当前帧中YOLOV5检测到的目标特征之间的最小余弦距离, 并对以上两种方法设定阈值, 设定匹配时的度量标准, 最终度量方式由二者的加权求和所得, 公式如下:

$$[0029] \quad c_{i,j} = \lambda d^{(1)}(i,j) + (1-\lambda) d^{(2)}(i,j) \quad (3)$$

[0030] 步骤3.4: 基于匈牙利算法对多个检测目标与跟踪轨迹进行匹配:

[0031] 使用匈牙利算法对DeepSORT跟踪器中的跟踪轨迹与YOLOV5当前所检测到的目标进行匹配, 利用步骤3.2中的特征提取模块, 以及步骤3.3中的距离度量标准组合成级联匹配, 对跟踪轨迹与检测目标之间进行距离计算, 并设定阈值过滤匹配结果, 之后还需要在IOU匹配中使用匈牙利算法对跟踪轨迹与检测目标再进行一次匹配, 最后利用匹配到的结果对当前所有轨迹分成三个状态确认态、未确认态、删除态。

[0032] 步骤4: 基于YOLOV5检测器与DeepSORT跟踪器结合进行飞机目标跟踪, 具体包括:

[0033] 从摄像头或者视频流中读入第一帧, 第一帧中使用YOLOV5检测器对帧中的飞机物体进行识别, 将所检测到的飞机目标的坐标信息、置信度以及飞机目标图像信息传递给DeepSORT跟踪器中, 并设置置信度阈值对检测框进行过滤, 删除置信度过小的检测框, 第一帧中无跟踪目标, 利用YOLOV5检测器对目标进行检测并在DeepSORT跟踪器中创建新的Tracker, 将检测到的目标存入tracks中, 并初始化卡尔曼滤波器状态, Tracker中的tracks状态初始化为未确认态。

[0034] 读入第二帧, YOLOV5检测后将检测信息输入到DeepSORT跟踪器中, 利用第一帧中创建的Tracker中的卡尔曼滤波器以及YOLOV5的坐标信息, 获得当前帧中的预测信息, 由于第一帧中无确认态的tracks, 则对未匹配的tracks以及未匹配的检测目标再进行IOU匹配, 计算出IOU最大的唯一匹配, 最后对Tracker中各个tracks的状态进行更新, 对于已匹配的tracks将所匹配目的特征保存到各个tracks的特征集中并更新卡尔曼滤波状态, 对于未匹配的检测目标创建新的tracks并进行初始化, 对于依然未匹配到的tracks直接删除。

[0035] 读入第三帧, YOLOV5检测后将检测信息输入到DeepSORT跟踪器中, 该帧中依然没有确认态的tracks, 在完成前后两帧检测目标与跟踪轨迹之间的IOU匹配后对已经连续三帧确认匹配的tracks变更为确定态, 对确定态的tracks进行特征更新, 并对卡尔曼滤波状态进行更新, 对于依然未匹配到的tracks直接删除, 对未匹配到的检测目标创建新的tracks并进行初始化。

[0036] 读入第四帧, YOLOV5检测后将检测信息输入到DeepSORT跟踪器中, 并使用步骤3.2中的特征提取模块对YOLOV5检测到的目标进行特征提取, 并利用tracks中的卡尔曼滤波信息以及特征信息, 以及步骤3.4中马氏距离、余弦距离以及匈牙利算法组合的级联匹配算法, 将处于确认态的tracks与当前帧中所提取到的目标特征进行匹配, 对已匹配的之前为确认态的tracks更新特征集以及卡尔曼滤波状态, 对已匹配的之前为未确认态的tracks匹配命中次数增加1, 对未匹配的tracks以及未匹配的检测目标再进行IOU匹配, 如果tracks仍未匹配且之前为确认态, 需要经过max_age帧后依然没有完成匹配, 则将tracks删除, 如果tracks仍未匹配且之前的状态为未确认态, 则直接删除tracks, 对未匹配的检测目标创建新的tracks并进行初始化, 在之后的每一帧中都与第四帧相同的方法进行匹配, 最后将跟踪框结果输出到视频中, 并将视频中的所跟踪到的飞机目标的坐标信息输出。

[0037] 优选地, 步骤3.3中的M和N都取值为70, 步骤4的置信度阈值取值为0.3。

[0038] 本发明的YOLOV5网络模型能够在密集场景下对飞机进行检测,并利用DeepSORT跟踪器进行实时跟踪。

[0039] 本发明具有以下优点:

[0040] 1.使用BiFPN多特征融合技术添加针对小目标的检测层,减少在检测阶段中对小目标的漏检问题。

[0041] 2.使用多重注意力机制CBAM模块对YOLOV5网络模型进行优化,增强YOLOV5检测的准确性。

附图说明

[0042] 图1本发明的Labelimg标注图片示意图。

[0043] 图2本发明的Darklabel标注视频示意图。

[0044] 图3本发明的基于BiFPN结构的多尺度特征融合结构图。

[0045] 图4本发明的改进型YOLOV5目标检测网络结构图。

[0046] 图5本发明的改进型YOLOV5目标检测模型训练结果评价图。

[0047] 图6本发明的YOLOV5检测器与DeepSORT跟踪器结合流程图。

[0048] 图7是本发明方法的流程图。

具体实施方式

[0049] 为了能够清晰地说明本发明地技术方法,下面结合附图对本发明进一步说明。

[0050] 本发明一种基于改进型YOLOV5算法的飞机多目标跟踪方法,步骤如下:

[0051] 步骤1:收集与标注飞机目标数据集,过程如下:

[0052] 收集飞机图像数据以及视频数据,使用标注工具Labelimg、Darklabel分别对图像和视频进行标注,并构建飞机图像数据集,将所述数据集划分成训练集、测试集以及验证集,其中使用Darklabel生成xml文件其中包含分类信息和坐标信息,并使用其所包含的分类和坐标信息将飞机目标从视频帧中裁剪出来,Labelimg首先生成包含分类信息和坐标信息的xml文件,再转换成符合YOLOV5训练格式的txt文件。

[0053] 步骤2:基于步骤1中的数据集对YOLOV5网络进行训练,并针对飞机目标进行改进,过程如下:

[0054] 步骤2.1:基于BiFPN结构进行多特征融合:

[0055] 在YOLOV5模型的head部分中使用BiFPN结构进行多特征融合,该结构使用高效的双向交叉连接进行特征融合,例如输入图片大小为640*640的图片,普通的YOLOV5网络会输出3个不同大小的特征图层80*80、40*40、20*20,为了检测小型飞机目标增加尺度为160*160更大的特征检测图层,BiFPN结构首先对于YOLOV5主干网络中获得的特征图进行Upsample上采样获得不同尺度的特征图,再使用步长为2的CONV卷积层进行下采样,并利用Concat操作在通道维度上结合主干网络中的特征图获得4种不同检测尺度的特征图。

[0056] 深层的特征图层携带有更强的语义特征,较弱的定位信息。而浅层的特征图层携带有较强的位置信息,和较弱的语义特征。通过BiFPN结构增强多个尺度上的语义表达以及定位能力。

[0057] 步骤2.2:使用CBAM注意力机制模块增加YOLOV5模型检测精准度:

[0058] 使用CBAM注意力机制,将通道注意力机制与空间注意力机制融合,通道注意力模块将特征图在空间维度上进行压缩,得到一个一维特征向量后再进行操作。将输入的特征图,分别经过基于宽和高的全局最大池化操作和全局平均池化操作,然后分别经过MLP。将MLP输出的特征进行基于按照对应元素逐个相加操作,再经过sigmoid激活操作,生成最终的通道注意力特征图。将该通道注意力特征图和输入的特征图按照对应元素逐个相乘。

[0059] 空间注意力模块将通道注意力模块输出的特征图作为本模块的输入特征图。首先做一个基于通道维度的全局最大池化操作和全局平均池化操作,然后将这两个结果基于通道维度做Concat操作。然后经过一个卷积操作,降维为1个通道。再经过sigmoid生成空间注意力特征图。最后将该特征图和该模块的输入特征图做按照对应元素进行相乘,最终得到生成的特征。该注意力机制所得特征能够覆盖到检测物体的更多部位,最终使判断物体的几率更高,增强模型的检测精准度。

[0060] 步骤2.3:获得YOLOV5检测模型:

[0061] YOLOV5网络包括主干网络backbone部分、head部分,主干网络backbone由CONV模块、C3模块、SPPF模块组成,其中CONV是由Conv2d卷积层、BN正则化层、Silu激活函数构成,C3模块是由3个CONV模块和多个Bottleneck模块并使用CSP结构所构成的模块,SPPF为金字塔结构的池化层,先使用使用1*1卷积核的CONV层并将通道数减半,再使用5*5的卷积核对已经经过CONV层的特征图进行三次最大池化操作,并依据次数获得3个特征图,最后将经过卷积层和池化操作所获得的4个特征图通过Concat操作在通道维度上进行拼接。

[0062] 首先对飞机数据集进行预处理,其中预处理包含Mosaic数据增强、K-means算法优化锚框以及自适应调节图像尺度,再将图像输入到主干网络backbone部分产生多种尺寸的特征图,然后head部分依照步骤2.1的BiFPN结构,进行Upsample上采样操作、在CONV中设置步长为2的下采样操作以及Concat按照通道进行交叉拼接操作完成多特征融合,再经过CBAM注意力机制模块产生4种不同尺度大小的特征图,将4个不同尺度的特征图输入head部分的Detect层中,能够得到所要检测目标的检测框位置、类别以及置信度,再利用多个损失函数结合进行反向传播,并对YOLOV5模型中的梯度进行更新调节权重参数,最后利用验证集对每个阶段的模型进行验证,获得YOLOV5检测模型。

[0063] 步骤3:构建DeepSORT跟踪器,过程如下:

[0064] 步骤3.1:构建DeepSORT跟踪器运动模型:

[0065] 使用卡尔曼滤波对已检测到的飞机目标构建运动模型并进行预测,预测下一帧中飞机目标位置,卡尔曼滤波适合用于解决线性问题,以及有预测结果的实时性,卡尔曼滤波可以只考虑当前帧下的飞机运动目标状态并预测下一帧中飞机的位置,更新时无需考虑之前帧中的检测值,只有上一帧中的预测值与当前帧中检测值相关。

[0066] 步骤3.2:构建DeepSORT跟踪器特征提取模块:

[0067] 使用步骤1中Darklabel对视频进行标注的图像数据,并对图像进行裁剪,对所得到的飞机图像数据进行分类,最后对数据集划分训练集和验证集。将所得数据输入到ResNeXt50网络中,并得到ResNeXt50特征提取模型,并使用该特征检测模型对YOLOV5所检测到的飞机目标进行特征提取。

[0068] 步骤3.3:针对DeepSORT跟踪器中目标前后帧跟踪轨迹与检测目标之间规定距离度量标准:

[0069] 首先使用马氏距离表示运动模型之间的关联度并作为距离函数：

$$[0070] \quad d^{(1)}(i, j) = (d_j - y_j)^T S_i^{-1} (d_j - y_j) \quad (1)$$

[0071] 其中 d_j 表示第 j 个YOLO V5检测到的目标位置, y_i 表示第 i 个DeepSORT跟踪器中所预测的目标位置, S_i^{-1} 表示跟踪到的目标位置与检测目标位置的协方差矩阵的逆矩阵, $d^{(1)}$ 表示在多维运动状态数据中YOLOV5所检测到目标的检测框与卡尔曼滤波所得到的预测框的关联程度。

[0072] 再加入特征之间的余弦距离,对检测到的目标特征与跟踪到的目标特征之间计算余弦距离：

$$[0073] \quad d^{(2)}(i, j) = \min \{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in \mathfrak{R}_i\} \quad (2)$$

[0074] 其中 r_j 表示第 j 个检测到所提取的目标特征并添加限制条件 $\|r_j\| = 1$ 所得, \mathfrak{R}_i 为每个跟踪目标构建一个轨迹,在其中存储每个跟踪目标在最近的70帧中关联成功的特征向量, $r_k^{(i)}$ 表示第 k 帧中第 i 个跟踪到的目标特征, $d^{(2)}$ 表示第 i 个跟踪到的最近70个成功关联的特征集与当前帧中YOLOV5检测到的目标特征之间的最小余弦距离,并对以上两种方法设定阈值,设定匹配时的度量标准,最终度量方式由二者的加权求和所得,公式如下：

$$[0075] \quad c_{i,j} = \lambda d^{(1)}(i, j) + (1 - \lambda) d^{(2)}(i, j) \quad (3)$$

[0076] 步骤3.4:基于匈牙利算法对多个检测目标与跟踪轨迹进行匹配：

[0077] 使用匈牙利算法对DeepSORT跟踪器中的跟踪轨迹与YOLOV5当前所检测到的目标进行匹配,利用步骤3.2中的特征提取模块,以及步骤3.3中的距离度量标准组合成级联匹配,对跟踪轨迹与检测目标之间进行距离计算,并设定阈值过滤匹配结果,之后还需要在IOU匹配中使用匈牙利算法对跟踪轨迹与检测目标再进行一次匹配,最后利用匹配到的结果对当前所有轨迹分成三个状态确认态、未确认态、删除态。

[0078] 步骤4:基于YOLOV5检测器与DeepSORT跟踪器结合进行飞机目标跟踪：

[0079] 从摄像头或者视频流中读入第一帧,第一帧中使用YOLOV5检测器对帧中的飞机物体进行识别,将所检测到的飞机目标的坐标信息、置信度以及飞机目标图像信息传递给DeepSORT跟踪器中,并设置0.3的置信度阈值对检测框进行过滤,删除置信度过小的检测框,第一帧中无跟踪目标,利用YOLOV5检测器对目标进行检测并在DeepSORT跟踪器中创建新的Tracker,将检测到的目标存入tracks中,并初始化卡尔曼滤波器状态,Tracker中的tracks状态初始化为未确认态。

[0080] 读入第二帧,YOLOV5检测后将检测信息输入到DeepSORT跟踪器中,利用第一帧中创建的Tracker中的卡尔曼滤波器以及YOLOV5的坐标信息,获得当前帧中的预测信息,由于第一帧中无确认态的tracks,则对未匹配的tracks以及未匹配的检测目标再进行IOU匹配,计算出IOU最大的唯一匹配,最后对Tracker中各个tracks的状态进行更新,对于已匹配的tracks将所匹配目标特征保存到各个tracks的特征集中并更新卡尔曼滤波状态,对于未匹配的检测目标创建新的tracks并进行初始化,对于依然未匹配到的tracks直接删除。

[0081] 读入第三帧,YOLOV5检测后将检测信息输入到DeepSORT跟踪器中,该帧中依然没有确认态的tracks,在完成前后两帧检测目标与跟踪轨迹之间的IOU匹配后对已经连续三帧确认匹配的tracks变更为确定态,对确定态的tracks进行特征更新,并对卡尔曼滤波状

态进行更新,对于依然未匹配到的tracks直接删除,对未匹配到的检测目标创建新的tracks并进行初始化。

[0082] 读入第四帧,YOLOV5检测后将检测信息输入到DeepSORT跟踪器中,并使用步骤3.2中的特征提取模块对YOLOV5检测到的目标进行特征提取,并利用tracks中的卡尔曼滤波信息以及特征信息,以及步骤3.4中马氏距离、余弦距离以及匈牙利算法组合的级联匹配算法,将处于确认态的tracks与当前帧中所提取到的目标特征进行匹配,对已匹配的之前为确认态的tracks更新特征集以及卡尔曼滤波状态,对已匹配的之前为未确认态的tracks匹配命中次数增加1,对未匹配的tracks以及未匹配的检测目标再进行IOU匹配,如果tracks仍未匹配且之前为确认态,需要经过max_age帧后依然没有完成匹配,则将tracks删除,如果tracks仍未匹配且之前的状态为未确认态,则直接删除tracks,对未匹配的检测目标创建新的tracks并进行初始化,在之后的每一帧中都与第四帧中相同的方法进行匹配,最后将跟踪框结果输出到视频中,并将视频中的所跟踪到的飞机目标的坐标信息输出。

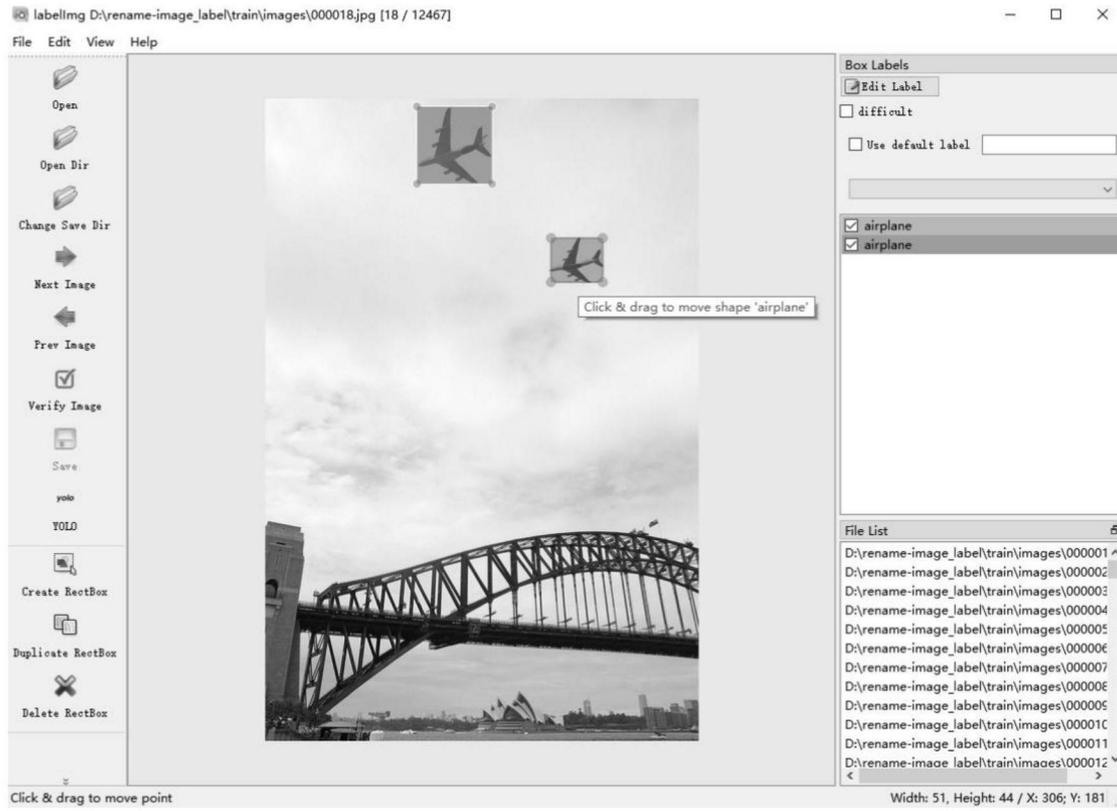


图1



图2

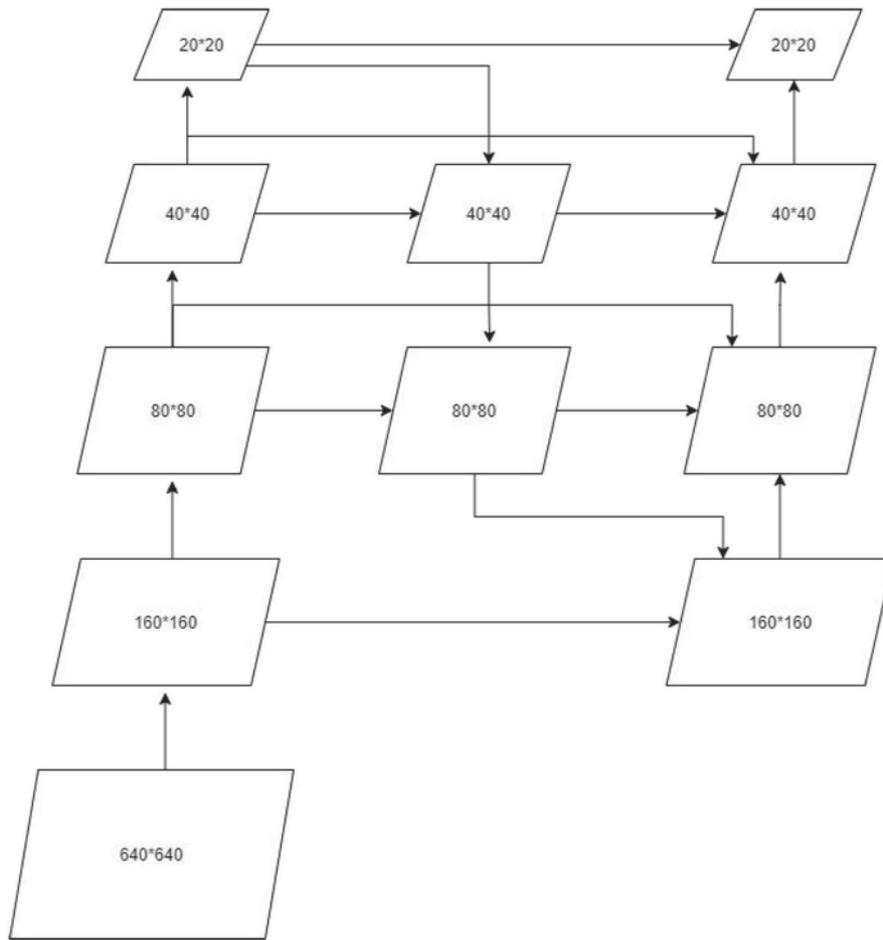


图3

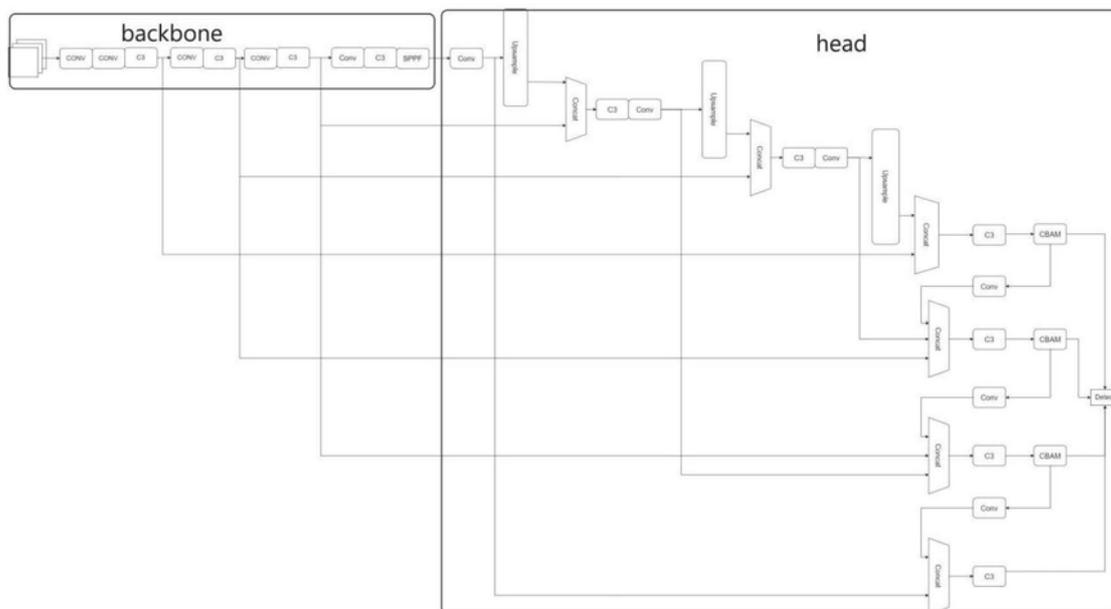


图4

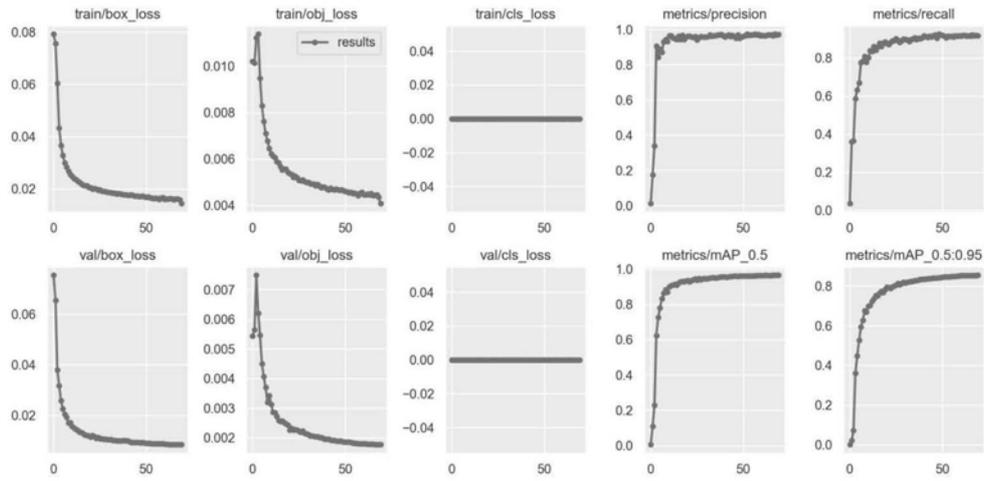


图5

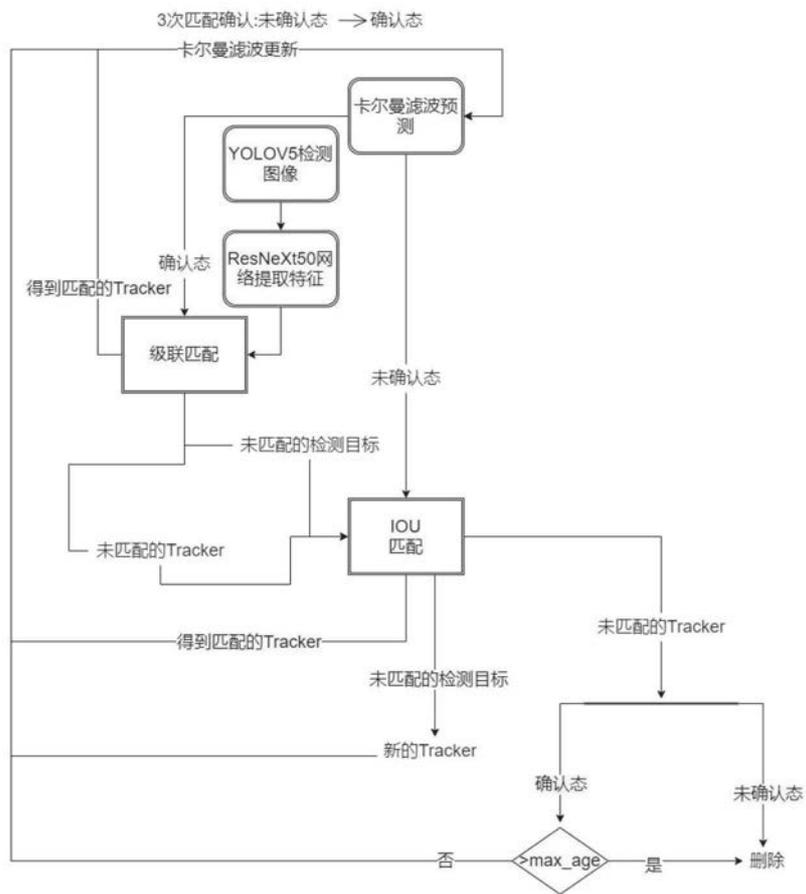


图6

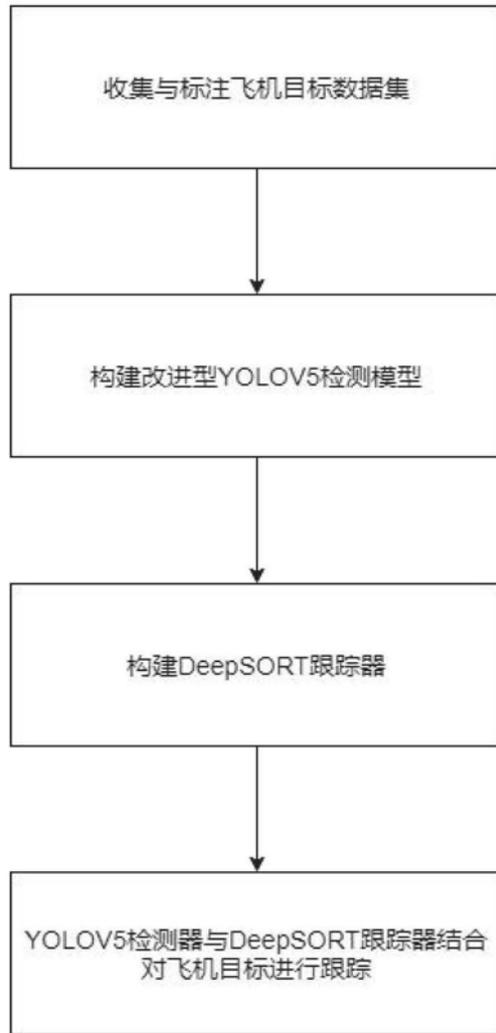


图7