

DOI: 10.13382/j.jemi.B2205301

基于改进的 YOLOv5 的交通锥标检测系统*

赵梓杉 桑海峰

(沈阳工业大学信息科学与工程学院 沈阳 110870)

摘要:针对中国大学生无人驾驶方程式赛车目标检测系统当前算法检测速度较慢,在不同场景下容易出现检测精度低,漏检、误检现象严重等问题,文章设计了一种适用于锥标颜色的整体检测系统,在识别模块中,首先,为了提高原 YOLOv5 基础模型检测速度和识别精度,采用 CIoU 作为边界框回归损失函数,针对训练时收敛速度慢和算法识别精确度低的问题,将原加权非极大抑制方式更改为 DIoU_NMS,测试精度为 0.963,相较于原算法提高了 2.1%,结果表明改进后的算法更适合比赛场景下锥标颜色识别。其次,在跟踪模块中,对深度表观特征锥标颜色重识别模型进行训练,将单目标跟踪算法改为可以对多种类别目标进行跟踪,相比于单一的目标检测算法,有效降低漏检现象,最后,添加测距模块,利用检测框高度信息进行车载摄像头到锥桶的测距,90 m 以内的平均误差小于 9%。整个系统的帧率达到 20 FPS,实现锥标颜色识别和距离的有效测量,为比赛提供更多的数据支持。

关键词:无人驾驶方程式赛车;交通锥标;深度学习;目标检测;目标跟踪;相似三角形测距

中图分类号: TP391;TN98 **文献标识码:** A **国家标准学科分类代码:** 520.20

Traffic cone detection system based on improved YOLOv5

Zhao Zishan Sang Haifeng

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

Abstract: In view of the slow detection speed of the current algorithm of the target detection system for Chinese college students' driverless formula racing cars, the low detection accuracy and serious missing and false detection in different scenarios are easy to occur. In the recognition module, first of all, in order to improve the detection speed and recognition accuracy of the original YOLOv5 basic model, the paper uses CIoU as the boundary box regression loss function. To solve the problems of slow convergence speed and low recognition accuracy of the algorithm during training, the original weighted nonmax suppression method is changed to DIoU_NMS in this paper, the test accuracy is 0.963, which is 2.1% higher than the original algorithm. The results show that the improved algorithm is more suitable for cone color recognition in the competition scene. Secondly, in the tracking module, the depth apparent feature cone color recognition model is trained, and the single target tracking algorithm is changed to be able to track multiple types of targets. Compared with a single target detection algorithm, the phenomenon of missing detection is effectively reduced. Finally, the ranging module is added to use the height information of the detection frame to distance the vehicle camera to the cone barrel. The average error within 90 meters is less than 9%. The frame rate of the whole system reaches 20 frames/second, realizing cone color recognition and effective distance measurement, and providing more data support for the game.

Keywords: driverless formula racing; traffic cone; deep learning; target detection; target tracking; similar triangle ranging

0 引言

无人驾驶技术中,感知层作为无人驾驶汽车的安全保障,在整车系统中占据了至关重要的位置^[1],能够为无人驾驶汽车提供实时的环境状况。

目前学者们将神经网络技术融入到无人驾驶环境感知当中,常用的传感器包括雷达和单目相机,单目相机具有价格低廉、能够对颜色进行识别等特点,是主要的发展趋势。

文献[2]提出基于交通标志的颜色和形状建立模板库。传统的基于机器学习的方法使用人工设计特征,然后用于训练分类器,最后再使用分类器判定类别。文献[3]中提出了一个两级卷积神经网络用于交通标志分类,SwinTransformer作为将Transformer应用于视觉检测任务的代表性算法,对于大目标的检测精度较高,但检测小目标的均值平均精度(mean average precision, mAP)仅31.6%^[4]。

环境感知技术也是赛车取得优异成绩的重要保证。其中,目标检测任务在赛车环境感知系统起着重要作用,包括对赛道和交通锥标的检测等。上述方法无法适用于无人驾驶方程式赛事特定场景。首先,无法增强对小目标的检测,其次,深层和复杂的网络使得计算量增加,最后对硬件的要求提高,会占用赛车布置空间的增加。

2018年,黄遥^[5]基于HSV设计出对大赛使用的交通锥标进行颜色识别的算法,但是在光照强烈情况下,检测系统会出现漏检或误检的情况,且该算法需要雷达将锥标的点云映射到对应的图像上,检测成本高。

张凯^[6]将深度学习引入到锥桶颜色识别工作中,使用价格低廉的相机作为感知系统的传感器,针对大赛锥标识别问题,选取YOLOv3^[7]算法,单帧识别耗时100ms,降低成本的同时置信度满足比赛要求,该算法还需进一步提高检测速度且存在一定程度的漏检和误检现象。

2020年,李涛等^[8]提出将基于改进的YOLOv3-tiny的深度学习框架用于大赛锥桶的颜色识别工作,该算法同时具有YOLOv3的检测精度,也提高了一定的检测速度,每秒识别的帧率达到20。

1 交通锥标检测系统

本文主要针对交通锥标感知系统进行研究。在无人驾驶方程式赛车行驶过程中,感知系统对前方交通锥标颜色进行识别和目标跟踪并进行实时测距,首先通过视觉传感器,对道路环境进行采集,对输入的视频帧进行交通锥标目标框提取,为了改善检测算法的漏检现象,在此基础上添加目标跟踪算法,最后利用检测框像素高度信

息对固定高度的交通锥标进行实时测距,将得到的交通锥标位置和距离信息反馈给无人驾驶方程式赛车的控制系统,以获得更优异的比赛成绩,避免不必要的比赛失误。其系统结构如图1所示。

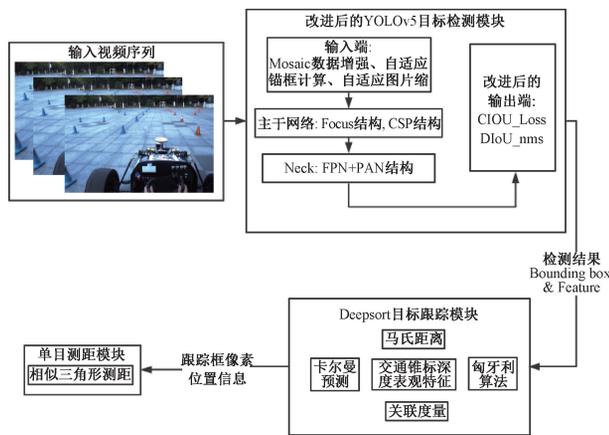


图1 交通锥标感知系统结构

Fig. 1 Structure of traffic cone perception system

1.1 改进的YOLOv5目标检测

YOLOv5网络主要包括4个部分:输入端、主干网络、Neck、输出端^[9],输出端主要由损失函数部分和非极大值抑制部分组成。

1) 损失函数改进

IoU为预测框和目标框的交并比(intersection over union, IoU),为了解决YOLOv3模型输出端中使用的损失函数IoU Loss不能优化预测框和目标框不相交情况的问题,并且不能反映两个框不相交时的情况。YOLOv5中的输出端使用了GIoU Loss,计算公式为:

$$GIoU_Loss = 1 - GIoU = 1 - IoU + \frac{|C/(A \cup B)|}{|C|} \quad (1)$$

IoU的取值在区间 $[-1, 1]$, $IoU = 0$, $GIoU = 1$ 在两个框无交集的情况下, $IoU = 0$, $GIoU = 1$,当两个矩形框完全重叠 $IoU = 1$, $GIoU = -1$,在非重叠的情况下,可以看到由于增加了惩罚,预测框向真值框移动。

GIoU同时关注重叠区域和非重叠区域,因此可以更好地反映预测框与真值框的重叠程度^[10],但当预测框与真值框对齐效果不佳时,则最小包围矩形C的面积增大,GIoU值减小^[11]。如果两个矩形不重叠,也可以计算GIoU值。当两个矩形属于包含关系时,GIoU会退化成IoU,丢失两个矩形的相对位置关系。

为了促使预测框向重叠区域增多的方向进行移动,进一步提升回归精度,本文边界框回归采用了CIoU Loss,CIoU Loss通过GIoU Loss得来,CIoU在DIoU的基础上将纵横比考虑到损失函数中,即在DIoU惩罚项的基

础上添加一个影响因子 av , 可以获得更加精确的位置信息。计算总的惩罚项:

$$R_{Clou} = \frac{\rho^2(b, b^{gt})}{c^2} + av \quad (2)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (3)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right) \quad (4)$$

$$CIoU_Loss = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (5)$$

其中, h^{gt}, w^{gt} 分别表示真值框高宽, h, w 分别表示预测框高宽, 平衡比例参数计算公式记为 α , 描述预测框和真值框长宽比一致性的参数记为 v 。当 $IoU=0$ 的时候, $CIoU$ 和 $DIoU$ 的优化方式是一样的, v 对 w 参数的偏导数公式如下:

$$\frac{\partial v}{\partial w} = \frac{8}{p^2} \left(\arctan \frac{w^{gt}}{h^{gt}} \arctan \frac{w}{h} \right) \times \frac{h}{w^2 + h^2} - \frac{8}{p^2} \left(\arctan \frac{w^{gt}}{h^{gt}} \arctan \frac{w}{h} \right) \times \frac{w}{w^2 + h^2} \quad (6)$$

本文选择的 $CIoU$ 方式计算边界框的回归损失, 可以更好地回归边界框的位置和大小。便于后续得到更加准确的得到检测框 4 个顶点的像素信息。

2) NMS 非极大值抑制改进

非极大值抑制 (non-maximum suppression, NMS) 即抑制不是最大值的元素, 一帧图像中感兴趣的物体往往是不同形状的, 其大小和宽高比都大不相同^[12]。因此目标检测算法往往会对同一个目标预测出不同的预测框, NMS 作为目标检测算法的最后一步, 用于剔除重复框, 然后仅保留某个检测框与预测框得分最高的交并比 (IoU) 的候选框, 并删除其他所有冗余框。

但是在目标密集的环境中, 由于各个物体之间存在遮挡, 不同目标的检测框离的非常近, 重叠面积比较大, 因此容易被非极大抑制错误地移除, 导致目标检测失败, 因此本文使用 $DIoU$ NMS 替换原始 NMS 来改善这个问题。如式(7)所示:

$$DIoU_Loss = 1 - DIoU = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} \quad (7)$$

其中, b, b^{gt} 分别代表预测框和真值框的中心点, ρ 用于计算两点间的欧氏距离 (两点之间的向量距离), d 表示预测框与真值框中心点的距离, 即 $d = (b, b^{gt})$, c 表示能够同时包含预测框与真值框的最小外接矩形的对角线长度。

$$s_i = \begin{cases} s_i, DIoU(M, B_i) < \varepsilon \\ 0, DIoU(M, B_i) \geq \varepsilon \end{cases} \quad (8)$$

通过对预测框与真值框之间的归一化距离进行建模, 直接最小化预测框与真值框的距离, 可以在预测框与

目标框不重合时为预测框提供移动方向, 从而加速模型的收敛。

M 表示预测分数最高的一个预测框, B_i 表示判断是否需要被移除的预测框, s_i 表示分类分数, ε 表示 NMS 的阈值。DIoU-NMS 会判断两个边界框 M 和 B_i 点之间的距离, 当距离较远时不会移除预测框, 而是认为检测到了另外的目标, 这有助于解决目标互相遮挡情况下的漏检问题。

1.2 Deepsort 交通锥标跟踪模块

在实际检测过程中, 外界光照不足, 相机抖动, 以及存在遮挡等复杂因素影响, 仅依赖检测结果, 会出现漏检的情况, 为了有效减少漏检, 故本文设计的感知系统增加了目标跟踪模块。

多目标跟踪任务的核心主要包括检测部分、预测部分、匹配部分和更新部分 4 个过程, 目标检测模块对视频相邻两帧图像感兴趣对象进行检测, 在此过程中, 提取特征向量采用裁剪对象的方式, 运动特征提取器对目标特征的相似度进行判断, 最后输出特征匹配程度, 最后为了让下一帧中匹配度高的目标分配前帧目标的身份识别码 ID, 进行数据关联^[13]。

1) 多目标跟踪流程

Deepsort 算法基于原有的 Sort 算法进行改进, 提取深度表观特征, 能更好地解决遮挡问题, 跟踪效果的优劣取决于现有的检测结果。跟踪流程如图 2 所示。

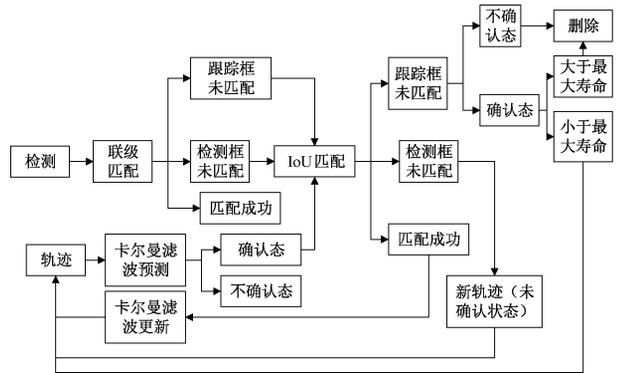


图 2 Deepsort 跟踪流程

Fig. 2 Deepsort tracking process

2) 深度表观特征

Deepsort 模型能实现对交通锥标良好的跟踪效果, 论文特提出交通锥标识别 C-RelD 模块, 在一定程度上提高了基于锥桶表观特征跟踪性能。

为优化跟踪模型适配性, 采用原深度表观特征卷积神经网络。以此网络对自建数据集进行预训练, 实际上, C-RelD 特征的训练采用了 1 个宽残差网络包括 2 个卷积层, 1 个最大池化层, 6 个残差块, 1 个密集层以及批处理和归一化层。其中密集层负责计算每个目标对应的 128

维度全局特征图的特征向量,也就是表现特征。

3) 计数显示类别

常见的目标跟踪算法不能同时跟踪多类别目标,大部分跟踪器只是应用在单类别多目标场景,仅在模型的训练过程添加多类别检测,无法实现多颜色锥标的跟踪,检测结果仅表示目标ID而无法标注类别。对Deepsort原算法代码进行编写,使其具备多类别的跟踪性能,跟踪时可识别出赛事所需的各种颜色锥标。

1.3 单目测距模块

目标检测模型的输出为图像中的像素与类别信息,在无人驾驶方程式赛车的感知系统中,这些信息远远不够,比如前方检测到了黄色锥桶,此时应该及时减速或制动,刹车的前提是需要知道交通锥标与赛车之间的距离信息,所以检测跟踪和测距是分不开的,这样的感知系统才有意义。

1) 相似三角形测距原理

在无人驾驶比赛场景下,由于比赛时使用的交通锥标都是同一尺寸,待检测物具有确定性,可以获取物体实际的高度,因此该方法适用于无人驾驶方程式赛车行驶过程中对于障碍物交通锥标的距离测量。小孔成像原理^[14]模型如图3所示。

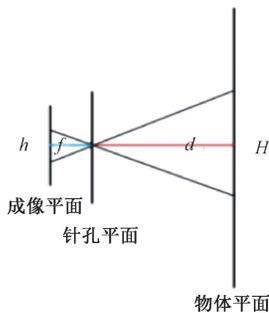


图3 测距模型图

Fig. 3 Ranging model diagram

为了获取物体的距离信息选择单目测距算法,目的是将单张图像从二维转换到三维。选用针孔成像原理建模,其中相机焦距为以像素为单位的 f 来表示,物体实际高度为 H ,物体在图像上的像素高度用 h 来表示,物体平面与相机平面的距离由 d 来表示,即所求距离。公式如下:

$$d = \frac{Hf}{h} = \frac{Hf}{(y - y_0)} \quad (9)$$

因此该方法适用于无人驾驶方程式赛车行驶过程中对于障碍物交通锥桶的距离测量。

2) 相机选型与安装

为了使基于深度学习算法的赛车在赛场上能够准确的识别出锥标的颜色信息,需要正确的选择单目相机的

型号,保证单目相机的性能可以使赛车适应不同的比赛环境。相机的参数主要包括分辨率、焦距和采集速率等。同时,这些参数也作为视觉传感器选型的主要因素。要保证识别算法的实时性,需要选择具有高分辨率以及高速采集效率的单目相机。除此之外,还需要考虑到比赛时的环境因素,因此,论文采用了大恒水星系列工业相机,满足赛车感知系统识别需求。该单目相机各项关键参数如表1所示。

表1 摄像头关键参数

Table 1 Key parameters of the camera

关键参数	参数值
分辨率	1 920×1 200
最小像素尺寸/ μm	4.4×4.4
帧率/FPS	30~60
工作温度/ $^{\circ}\text{C}$	-30~45
工作湿度/%	10~80
工作电压/V	9~12

所选的大恒水星单目相机,镜头较长且整体突出,安装位置如图4所示。由比赛规则可知,车体所搭载的传感器安装位置不能干扰到赛车手的操作,因此,为了确保单目相机的可视范围,将单目相机安装在无人驾驶方程式赛车座椅的上方。



图4 摄像头在车上的安装位置

Fig. 4 The installation position of the camera on the car

赛车过弯时靠近车体附近的锥标处于单目相机的视野盲区,尽管赛车可以依据激光雷达提供的锥标位置,做出规划控制,但是图像颜色数据的缺失同样会引起赛车错误的运动。通过使用单目相机在转弯等特殊运动状态下可观察到车体附近锥标。如图4所示,使用一对直角块固定相机,防止因相机位移造成标定位置产生误差。

2 实验结果与分析

2.1 数据集

本文的研究内容是建立无人驾驶方程式赛车的实时感知系统,其中检测模块的检测对象为构成赛道的3种交通锥标,分为红色锥桶、蓝色锥桶、黄色锥桶,有关比赛场景下的交通锥标图片以及标注较为匮乏。因此本文自制了交通锥标颜色检测数据集。

本文的标注的对象分别为红色锥桶 (red)、蓝色锥桶 (blue)、黄色锥桶 (yellow)。使用 LabelImg 开源工具对数据集中的采集的样本进行标注^[15],标注过程界面如图 5 所示。

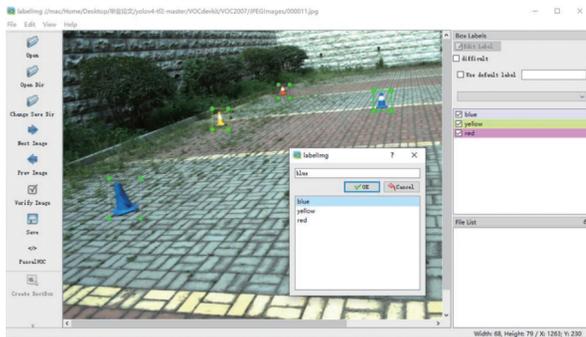


图 5 数据集标注实例

Fig. 5 An example of dataset annotation

2.2 相机内参

论文共拍摄了 46 张棋盘平面图像,其中 23 张为近距离拍摄(A 组),23 张为远距离拍摄(B 组)。相机的内部参数如表 2 所示,由 MATLAB 工具箱进行标定得到。

表 2 相机内部参数

Table 2 Camera internal parameters

参数	A 组	B 组	A+B 组
f_x	177.461 9	200.187 3	199.933 6
f_y	177.015 4	200.357 8	199.824 6
c_x	170.042 7	171.258 1	171.203 9
c_y	128.973 7	129.688 3	128.746 9
径向畸变	(-)	(-)	(-0.086 6, 0.083 1)
平均投影误差	0.088 9	0.092 1	0.091 1

2.3 评价指标

评价指标的目的是可以显示算法在不同性能上的相关体现。

1) 帧率 (frame per second, FPS) 即每秒检测视频帧数,每秒帧率越高,实时性越好。

2) 通过标注数据,可以将真实情况下的正样本,负样本分好类,并从上述过程中获得 4 个基本指标^[15] 如表 3 所示。

表 3 混淆矩阵

Table 3 Confusion matrix

真实样本	预测为正	预测为负
正样本数	TP	FN
负样本数	FP	TN

根据以上 4 个内容定义常用的目标检测评价指标。

3) 召回率 (Recall)

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (10)$$

式中: FN 为被检测出分类正确的锥标数, TP 为没有被检测出的锥标数, FP 为被检测出且分类错误的锥标数。

4) 精确度 (Precision)

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (11)$$

5) 准确率 (Accuracy)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

6) 平均精确度 (average precision), 是目标检测领域中最常用的评价模型性能指标,它的取值范围是 0~1。根据式 (13) 求出 AP:

$$AP = \frac{1}{11} \sum_{r=(0,0,L,1,0)} AP_r \quad (13)$$

7) 平均精确度均值 (mAP) 是 AP 的均值,当数据集中为单个目标时,mAP 和 AP 的含义相同,当数据集中含有多个种类时,mAP 即为每个样本 AP 的平均值^[16]。

2.4 实验分析

本章实验在 Ubuntu18.04 操作系统上,主板型号为微星 B450MMORTAR MAX (MS-7B89), GPU 为 Nvidia GeForce GTX1650super,可使用内存 16 GB,使用 CUDA 版本为 11.2,超参数中的权重衰减系数为 0.003 2;Batch 大小设置为 4;学习率动量为 0.843;迭代次数为 300。

1) 检测模块实验

(1) 定性分析

比赛场景下,当光照较弱,远处目标太小等条件下往往会导致识别算法失效,如图 6、7 所示。



图 6 传统 HSV 算法在弱光条件下的识别效果

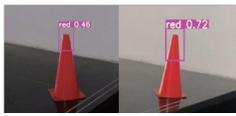
Fig. 6 Recognition effect of traditional HSV algorithm under weak light conditions



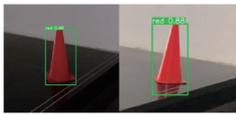
图 7 YOLOv3 远处小目标的识别效果

Fig. 7 Recognition effect of YOLOv3 remote small target

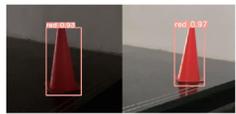
为了验证本文方法在恶劣环境下的鲁棒性,故在不同光强,不同锥桶尺寸,不同锥桶视角下、遮挡前后的锥桶和多目标锥桶颜色识别情况下对训练好的 3 个模型进行了定性实验。实验结果如图 8~12 所示(其中 a 为原车使用的 YOLOv3 模型,b 为原 YOLOv5 模型,c 为改进后的 YOLOv5 模型),图上的数值为置信度。



(a) YOLOv3 的检测效果
(a) Detection result of YOLOv3



(b) YOLOv5 的检测效果
(b) Detection result of YOLOv5



(c) 改进 YOLOv5 的检测效果
(c) Detection result of improved YOLOv5

图 8 不同光照情况下的检测效果对比
Fig. 8 Comparison of detection effects under different lighting conditions



(a) YOLOv3 的检测效果
(a) Detection result of YOLOv3



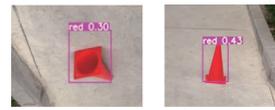
(b) YOLOv5 的检测效果
(b) Detection result of YOLOv5



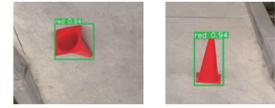
(c) 改进后 YOLOv5 的检测效果
(c) Detection result of improved YOLOv5

图 9 不同尺寸锥桶检测效果对比
Fig. 9 Comparison of detection effects of cone barrels of different sizes

本文使用数据集大小为 1 000 张,包括遮挡、未遮



(a) YOLOv3 的检测效果
(a) Detection result of YOLOv3



(b) YOLOv5 的检测效果
(b) Detection result of YOLOv5



(c) 改进后 YOLOv5 的检测效果
(c) Detection result of improved YOLOv5

图 10 不同锥桶视角的检测效果对比
Fig. 10 Comparison of detection effects of different cone-barrel viewing angles



(a) YOLOv3 的检测效果
(a) Detection result of YOLOv3



(b) YOLOv5 的检测效果
(b) Detection result of YOLOv5



(c) 改进后 YOLOv5 的检测效果
(c) Detection result of improved YOLOv5

图 11 带遮挡情况下的检测效果对比
Fig. 11 Comparison of detection effects with occlusion

挡、光线不足、光线充足、摆放角度、大小锥标共 3 139 个样本。从实验结果可以看出,在不同条件下,都能检测到交通锥标,但 YOLOv3 的 anchor 框有一定概率出现回归不准确的情况。且改进后的 YOLOv5 的置信度整体明显高于原 YOLOv5 模型,从图 12 可以看出,3 种模型都能对多锥桶的情况进行颜色识别,YOLOv3 中用于检测交通锥桶颜色的特征层提取的语义信息不够丰富,较低的语义信息会导致一定的分类错误或较低置信度,给颜色识别带来误检和漏检,如图 12(a)所示,出现了红黄锥桶误检以及红色锥桶漏检的现象,而本文算法的训练权重是

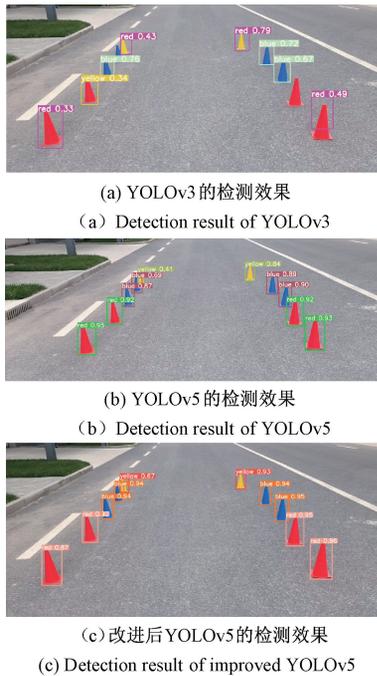


图 12 多锥桶颜色识别效果对比
Fig. 12 Comparison of color recognition effect of multi-cone barrel

在 YOLOv5 的预训练权重上进行训练的 实际上识别对象的特征(颜色,形状)相对来说比较简单 在预训练权重上进行微调达到比较好的效果在少量样本情况下仍能 达到 0.83。

(2) 定量分析

将 HSV、YOLOv3、YOLOv3-tiny、YOLOv5^[17] 和改进后的 YOLOv5 模型进行对比实验,光线弱和有遮挡的情况是影响比赛中锥标检测效果的主要因素,如图 6、7 所示,因此重新制作测试集,进行定量评估来验证改进后 YOLOv5 的有效性,实验结果如表 4 所示。

由表 4 可见,HSV 算法在恶劣条件下识别失效,且帧率过低,上述 4 种基于深度学习模型都可以满足实时性,改进之后的 YOLOv5 算法相较于原车所使用 YOLOv3 的测试精度提高了 32.1%,在满足实时性的条件下,改进的 YOLOv5 模型上均值平均精度(mAP)值、召回率(Recall)高于原 YOLOv5 模型,综合考虑,改进的 YOLOv5 模型在交通锥标颜色识别上取得一定的效果,且在恶劣比赛场景下,仍能保持最优的均值平均精度。

2) 跟踪模块实验

本文使用的改进后的 YOLOv5 检测网络输出为:图像帧数、对象的置信度、对象类别、检测框像素信息,在

表 4 网络对比实验结果

Table 4 Network comparison experimental results

模型	测试精度	mAP	小目标场景下的 mAP	弱光条件下的 mAP	Recall	训练时间/h	帧数/s	参数量/M
HSV	0.357	0.378	0.160	0.192	0.119	9.940	1.667	355.4
YOLOv3	0.621	0.521	0.377	0.395	0.218	2.889	26.081	492.7
YOLOv3-tiny	0.712	0.932	0.679	0.669	0.941	1.675	20.370	50.2
YOLOv5	0.942	0.973	0.758	0.707	0.959	0.663	32.298	14.4
改进后 YOLOv5	0.963	0.977	0.793	0.732	0.960	9.570	33.598	333.3

实际比赛中,检测环境受到光照、遮挡、相机抖动等多因素影响,所以仅依赖检测的结果,检测框存在不稳定的情况,下面展示仅有检测和与跟踪模块结合的检测结果对比。

由图 13 所示,加入跟踪模块处理的检测框更加稳定,输出结果更加平滑,还给出了不同对象的标签和同一类对象的不同 ID 信息,即对不同物体做了区分,又对同一类型物体做了区分,即实现了多目标跟踪,且将同一帧检测模块没有检测出来的锥桶检测了出来,有效降低了漏检率。

3) 测距模块实验

实验数据通过车载单目相机采集,在赛车前方每 15 m 放置一个锥标,测距上限为 90 m,由比赛规则可知,两端都为弯道的赛道长度不超过 60 m,一端为弯道的赛道长度不超过 45 m,通过此过程测试距离算法有效性。

由于在行驶过程中会导致检测框不稳定,所以会导致测量距离数据有一定范围内的波动,因此取每个位置

的误差平均值为最终误差,公式如下:

$$\text{平均误差} = \frac{1}{k} \sum_f \frac{|pd_f - gt_f|}{gt_f} \times 100\% \quad (14)$$

其中, k 是视频帧数, pd_f 表示预测值, gt_f 表示真实距离。将每个位置的每帧误差累加再求均值得到最终误差,实验结果如表 5 所列。

表 5 相似三角形测距结果

Table 5 Ranging results of similar triangles

模型	实际距离/m	测量距离/m	误差/%
YOLOv5+Deepsort+相似三角形测距算法	15	16.57	10.47
	30	32.66	8.86
	45	46.65	3.67
	60	65.12	8.53
	75	67.65	9.8
	90	80.13	10.97
平均误差/%		8.716	
FPS		20	



图 13 目标检测与加入跟踪处理的检测模块的检测结果对比

Fig. 13 Comparison of detection results between target detection and detection modules with added tracking processing

为了实时验证锥标检测系统的性能,采集 30 ~ 60 FPS 的车载单目相机视频流,将视频流分割为视频帧输入交通锥标检测系统。

此外,车辆在赛道上的行驶速度一般控制在 20 ~ 25 km/h 之间。如果以速度上限 25 km/h 计算,抓拍得到图像,测距系统处理此帧图像期间,车辆仍会继续前进,行驶距离为 0.185 m 左右,因此测距功能对安全驾驶有重要意义。其部分实验结果如图 14 所示,有效提供了锥桶种类,锥桶数量,锥桶与车的距离相关信息,为下游的规划模块和整车控制提供了有效的数据支持。

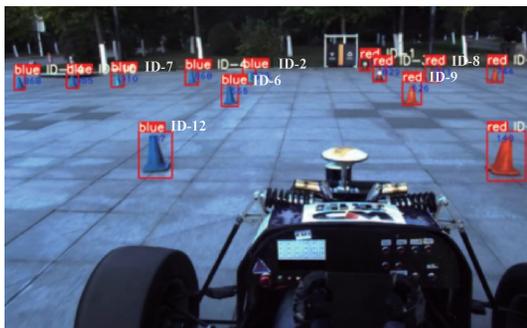


图 14 锥标检测+跟踪+测距实验示例

Fig. 14 Example of cone mark detection + tracking + ranging experiment

3 结 论

针对不同比赛场景下,锥标颜色检测速度慢,漏检、误检现象严重等问题,本文设计了一种适用于锥标颜色的整体检测系统,论文通过 CIoU 的方式计算边界框的回归损失,更好地回归边界框的位置和大小,便于后续得到

更加准确的得到检测框 4 个顶点的像素信息,并将原始 NMS 改进为 DIoU_NMS,通过该方法可以有效解决在目标相互遮挡时的漏检问题,其次论文基于 Deepsort 多目标跟踪算法,融合改进后的检测模型,进一步改善了单一目标检测任务中容易出现的漏检现象。最后添加测距模块,为后续路径规划和整车控制提供了更多的数据支持,整套算法在满足实时性的基础上有效提高了测试精度和合理性,未来的研究方向是采用压缩剪枝等手段使模型更加轻量化,降低工控机的算力需求。

参考文献

[1] 王小龙. 基于深度学习的目标检测与跟踪技术研究及应用[D]. 武汉: 湖北工业大学, 2020.
WANG X L. Research and application of target detection and tracking technology based on deep learning [D]. Wuhan; Hubei University of Technology, 2020.

[2] GAO X W, PODLADCHIKOVA L, SHAPOSHNIKOV D. et al. Recognition of traffic signs based on their colour and shape features extracted using human vision models[J]. Journal of Visual Communication and Image Representation, 2006, 17(4) : 675-685.

[3] MAO X, HIJAZI S, CASAS R, et al. Hierarchical CNN for traffic sign recognition [C]. 2016 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2016: 130-135.

[4] LEE C Y, CHIEN H C, WANG C P, et al. Using multi-scale SwinTransformer-HTC with data augmentation in CoNIC challenge[J]. arXiv preprint arXiv:2202.13588, 2022.

[5] 黄遥. FSAC 赛车雷达障碍物检测与摄像头图像识别研究[D]. 锦州: 辽宁工业大学, 2019.
HUANG Y. Research on FSAC racing radar obstacle

- detection and camera image recognition [D]. Jinzhou: Liaoning University of Technology, 2019.
- [6] 张凯. 无人驾驶电动方程式赛车环境感知算法研究[D]. 锦州: 辽宁工业大学, 2021.
ZHANG K. Research on environment perception algorithm of driverless electric equation car [D] Jinzhou: Liaoning University of Technology, 2021.
- [7] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.
- [8] 李涛, 张代胜, 张飞, 等. 基于改进 YOLO 算法的 FSAC 赛车实时目标检测系统[J]. 合肥工业大学学报(自然科学版), 2020, 43(9): 1153-1158.
LI T, ZHANG D SH, ZHANG F, et al. FSAC racing real time target detection system based on improved YOLO algorithm [J] Journal of Hefei University of Technology (Natural Science Edition), 2020, 43(9): 1153-1158.
- [9] 王莉, 何牧天, 徐硕, 等. 基于 YOLOv5s 网络的垃圾分类和检测[J]. 包装工程, 2021, 42(8): 50-56.
WANG L, HE M T, XU SH, et al. Waste classification and detection based on YOLOv5s network [J]. Packaging Engineering, 2021, 42(8): 50-56.
- [10] 伍冯洁, 陈江涛, 吕明瑞, 等. 一种跨场景目标自动识别与追踪方法及应用; , CN112801018A [P]. 2021.
WU F J, CHEN J T, LYU M R, et al. A cross scene target automatic recognition and tracking method and application; , CN112801018A [P]. 2021.
- [11] 龙燕, 李南南, 高研, 等. 基于改进 FCOS 网络的自然环境苹果检测[J]. 农业工程学报, 2021, 37(12): 307-313.
LONG Y, LI N N, GAO Y, et al. Apple detection in natural environment based on improved FCOS network [J]. Journal of Agricultural Engineering, 2021, 37(12): 307-313.
- [12] 尹昱航. 基于特征融合的交通场景目标检测方法研究[D]. 大连: 大连理工大学, 2021.
YIN Y H. Research on traffic scene target detection method based on feature fusion [D]. Dalian: Dalian University of Technology, 2021.
- [13] 许伟, 熊卫华, 姚杰, 等. 基于改进 YOLOv3 算法在垃圾检测上的应用[J]. 光电子·激光, 2020, 31(9): 928-938.
XU W, XIONG W H, YAO J, et al. Application of improved YOLOv3 algorithm in garbage detection [J]. Optoelectronics·Laser, 2020, 31(9): 928-938.
- [14] 徐守坤, 倪楚涵, 吉晨晨, 等. 基于 YOLOv3 的施工作业安全帽佩戴的图像描述[J]. 计算机科学, 2020, 47(8): 233-240.
XU SH K, NI CH H, JI CH CH, et al. Image description of safety helmet wearing in construction scene based on YOLOv3 [J] Computer Science, 2020, 47(8): 233-240.
- [15] 黄剑华, 丁建睿, 刘家锋, 等. 基于局部加权的 Citation-KNN 算法[J]. 电子与信息学报, 2013, 35(3): 627-632.
HUANG J H, DING J R, LIU J F, et al. Citation-KNN algorithm based on local weighting [J]. Journal of Electronics and Information Science, 2013, 35(3): 627-632.
- [16] 高照. 基于深度学习的遥感图像目标检测方法研究[D]. 青岛: 山东科技大学, 2020.
GAO ZH. Research on remote sensing image target detection method based on depth learning [D]. Qingdao: Shandong University of Science and Technology, 2020.
- [17] JOCHER G. YOLOv5[EB/OL]. <https://github.com/ultralytics/YOLOv5>, 2020.

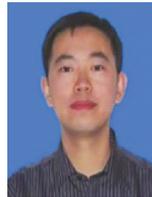
作者简介



赵梓杉, 2019 年于辽宁工业大学获得学士学位, 现为沈阳工业大学信息科学与工程学院博士研究生, 主要研究方向为目标检测、车辆轨迹预测。

E-mail: 1299113812@qq.com

Zhao Zishan received her B. Sc. degree from Liaoning University of Technology in 2019. Now she is a Ph. D. candidate at the School of Information Science and Engineering, Shenyang University of Technology. Her main research interests include target detection and vehicle trajectory prediction.



桑海峰(通信作者), 沈阳工业大学信息科学与工程学院教授, 主要研究方向为机器视觉检测、生物特征识别、深度学习、行人轨迹预测。

E-mail: sanghaif@163.com

Sang Haifeng (Corresponding author) is a professor at the School of Information Science and Engineering, Shenyang University of Technology. His main research interests include machine vision, biometrics identification, deep learning, pedestrian trajectory prediction.