****

**电气与电气工程学院**

**2022-2023学年第一学期**

**招商证券人工智能工程训练营**

**综合作业报告**

|  |  |  |
| --- | --- | --- |
| 班级 | 学号 | 姓名 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

华中科技大学电气与电子工程学院

2022 年9月

# 课程项目设计报告

1 研究背景和项目目标

**1.1选题依据**

物体识别与跟踪在现实世界中应用广泛，涉及智能交通系统，先进辅助驾驶，导弹制导，医疗诊断，视频监控，游戏等领域。该题目可以利用两种优势算法更好地实现物体识别与追踪，也就可以被更好地利用在各个领域。

**1.2业界现状介绍**

感觉是人类与外界联系的窗口和交流的桥梁，然而，由于人类的精力毕竟是有限的，人类的视野也是有限的，所以人类的视觉在各种领域的应用都受到很大限制甚至是低效的。计算机视觉研究的目的是使计算机感知环境中的物体的几何信息，包括它的形状、位置、姿态、运动等，并对其进行描述、存储、识别与理解，因此成为当今最热门的课题之一。目前，目标跟踪技术已经被广泛应用于众多生活和工作领域。而在这个系统中，其核心部分就是运动目标检测与跟踪算法技术，这个技术是保证整个系统实现其功能的关键。

目标跟踪是视觉领域研究的热点问题，随着计算机技术的发展而得到迅猛的发展，目标跟踪技术也由此取得了长足的进步。上世纪对图像的处理主要是集中于单幅图像的处理。直到上世纪80年代HomBK等人提出了光流法。在智能监控方面，Maryland大学研发的实时监控系统W4实现了人的跟踪，可用于对人的行为监视，并可以判断人是否携带物体等简单行为。1975年Fukunaga等人在一篇关于概率密度梯度函数的估计中首次提出Mean shift（均值偏移）概念。用mean-shift算法对目标跟踪有着速度快，具有较强的抗干扰能力。但是对运动中发生了形变、缩放、遮挡等变化的目标进行长时间跟踪也存在问题。

目标跟踪即使主要有两种实现方式，分别是基于检测的方法和基于识别的方法。

基于检测的方法一般是利用目标的特征直接在图像序列中提取目标。

基于识别的方法也称为基于匹配的方法。这种方法与基于检测的方法不同之处在于，它不要求得到背景（或非目标）部分的描述，不需要描述目标和背景的区别，而仅仅需准确描述目标自身的特征。

**1.3 本项目的目标**

目标识别与跟踪技术在生产生活的各个领域都有着重要的应用，本项目针对该技术在视频监控和智能交通领域的应用，运用YOLOv5和DeepSort技术实现人车等物体的的识别与跟踪，具体功能如下：

①基于YOLOv5和DeepSort实现视频和摄像头内人、自行车、轿车、摩托车、公交车、卡车等物体识别与跟踪，并且对其进行编号；

②设计GUI界面，实时展示视频和摄像头中物体识别与跟踪的画面；

③能够在GUI界面中框选某个对象，并跟踪被选定的对象；

④能够跟踪框选的对象并对其进行放大，使得目标对象始终在画面中央。

2 项目总体设计

该项目主要包括用户选择待处理的视频文件、视频中目标的识别与标记、视频中目标的跟踪、设计界面让用户框选目标、框选目标的放大五个模块。

**2.1 文件选择**

该模块主要实现的功能是通过交互式的界面，让用户选择所需要的文件界面，操作直观简单，方便用于使用，实现的复杂度不高，通过编写函数getfilepath即可实现。

**2.2目标的识别与标记**

该模块主要实现的目标是对视频中的人、自行车、轿车、摩托车、公交车、卡车等物体进行识别，并寻找对应的标签进行标记和编号。由于YOLOv5模型的训练需要大量的视频数据，训练速度较慢，训练花费时间过长，自己进行训练的难度较高，本项目采用已经训练完成YOLOv5模型，对其进行进一步运用，通过编写detector函数实现。

**2.3目标的跟踪**

该模块主要实现的功能是对目标的跟踪，运用deepsort算法可以实现，通过参考Github资源编写tracker实现，技术实现难度不高。

**2.4单个目标的框选**

该模块实现的功能时当用户按下‘0’时截取该帧图片，待用户框选一个矩形区域后跟踪该区域内的目标，基于YOLOv5与deepsort之前的模块，实现该功能复杂度不高，通过调用cv2库selectROI即可实现。

**2.5目标的放大**

该模块主要实现的功能是对所框选的目标进行插值放大，由于python中图像是以数组存储的，故只需要读取用户框选矩形框的四个点坐标，对视频每一帧的图像进行截取，再调用cv2库的resize函数对图像进行放大即可实现，实现的复杂度不高。

**2.6项目整体组织**

项目的各个模块能够实现既定的功能后对各个模块进行整合，最终的程序能够实现运行程序，用户通过用户界面选择待处理的视频文件即可进入识别模式。运行时按‘0’进入选择模式，框选一个目标，按Enter即可进行单目标跟踪，按‘2’进入放大模式，按‘3’退出放大模式，按ESC关闭程序。整个程序功能完整。

3 项目关键技术

**3.1 语言与工具**

本项目包含的文件较多，VSCode软件最大的优势是能够yi文件夹形式同时打开多个文件，方便快捷。除此之外，VSCode还拥有丰富的插件系统，可以编辑HTML、CSS、JS、TS、Vue、React等前端代码和JAVA、Python等后端代码。VSCode具有代码跟踪功能，可以对代码进行标注，可以查看更改了那些文件，文件中修改了第几行的代码等。鉴于VSCode的种种优点，我们选用其作为本项目的编程工具。

Python语言具有开发效率高的优点，Python具有非常强大的第三方库，在此基础上进行开发，很大程度上提高开发效率；同时Python语言面向对象python的编程思想更符合人类的思维逻辑，不需要过多的理解计算机的运行思维，Python语言再机器学习领域有着得天独厚的优势，故选用Python作为该项目的编程语言。

**3.2 模型与算法**

①基于YOLOv5的目标识别模型

本项目的核心模型是通过YOLOv5训练出来的物体识别模型。使用Pytorch框架的YOLOV5有非常轻量级的模型大小，YOLOV5的训练阶段使用传统的填充方式，即将图片统一缩放到标准的尺寸，但在测试阶段使用缩减填充的方式进行填充以提高速度，该缩减填充的方法是基于减少冗余信息影响的目的提出的。YOLOV5在测试阶段自适应的补充以实现最少的冗余信息， YOLOV5的结构借鉴了CSPNet的设计思路，在主干网络中设计了CSP结构，并且在网络的不同部位使用不同结构的CSP，从而降低计算和内存成本。因此YOLOv5

是一个在实现复杂度和成本上都不太难的目标识别算法，因此我们采用基于YOLOv5的目标识别模型作为本项目的模型。

②Deepsort算法

Deepsort多目标跟踪算法便是这类算法中较为常用的一个。Deep-sort 是在 sort 目标跟踪的基础上引入了表观特征提取网络，该网络是在行人重识别数据集上离线训练得到的，在目标跟踪过程中提取目标的表观特征进行最近邻匹配，这样可以有效的改善遮挡情况下的目标追踪效果。算法可以分为四个核心部分，轨迹处理和状态估计、相关性度量、级联匹配、深度特征描述器。轨迹处理和状态估计部分的算法和 sort 算法类似，主要分为运动状态估计和目标创建与移除。Deepsort 算法通过(u,v,gama,h,x1,y1,gama1,h1), 这八个参数描述目标的运动状态，其中 (u,v) 为目标锚点框的中心坐标，gama是目标锚点框的长宽比，h 表示目标锚点框的高度，其余四个变量为目标对应在图像坐标种的运动信息。通过一个基于常量速度模型和线性观测模型的标准卡尔曼滤波器对目标的运动状态进行预测，预测的结果为(u,v,gama,h)。设置一个阈值Amax，设置一个计数器ak 表示第 k 个追踪目标最后一次出现到当前帧所间隔的帧数，每次目标出现都将刷新计数器，即将计数器置为 0，若ak>Amax则，目标 k 追踪结束。若目标匹配过程中，有目标始终无法与已存在路径相匹配，则认为该目标可能为新出现的目标，如果接下来的 3 帧都连续检测到了该目标，则该目标将被认定为新出现的跟踪目标，并以此目标为起始目标生成新的跟踪路径，否则不生成新的跟踪路径。其运行流程图如下



**3.3 其他库**

本项目中使用的最多的是OpenCV库，OpenCV跨平台计算机视觉库， 在计算机视觉项目的开发中，OpenCV作为较大众的开源库，拥有了丰富的常用图像处理函数库，具有 Python等的使用接口、机器学习的基础算法调用，在图像处理领域有着重要应用。本项目采用OpenCV库对图像进行各种处理。

4 项目实现

**4.1 文件选择的实现**

文件选择通过编写getfilepath函数实现，函数中调用TK接口，通过filedialog.askopenfilename()即可实现用户界面文件路径的获取，再通过调用OpenCV库的cv2.VideoCapture(file\_path)函数即可打开视频文件。

**4.2目标的识别与标记的实现**

在Github中寻找训练好的YOLOv5目标识别模型，进行测试，最终筛选出合适的模型作为本项目使用的模型。

**4.3目标的跟踪的实现**

在Github上搜索Deepsort算法实现范例，再结合YOLOv5模型，编写函数tracker。先在 YOLOv5 的 bbox 中搜索中心点最接近的label，然后计算距离阈值，判断是否在允许误差范围内。若满足误差则进行赋值工作，首先判断是否是第一次赋值，若是则取 x, y 方向上的距离平均值进行赋值，若不是则取原值与x, y 方向上的距离平均值进行赋值，由此循环进行目标的跟踪。

**4.4单个目标的框选的实现**

通过调用OPenCV库selectROI即可实现对目标的选择与跟踪。OpenCV库中自带的KCF跟踪函数可以实现对包括但不限于人和车的物体的跟踪。

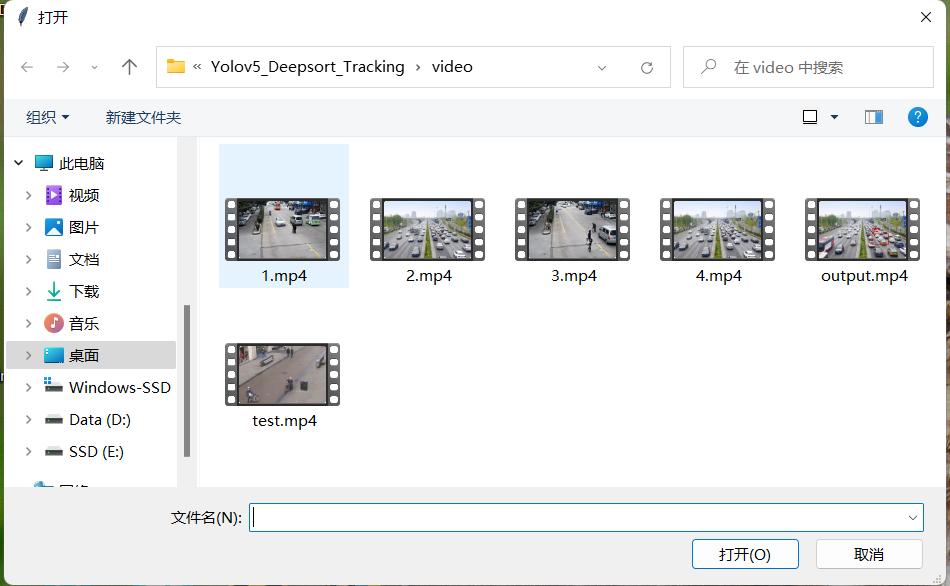
**4.5目标的放大的实现**

通过获取用户所框选的矩形框四个的的坐标，对每一帧图片进行截取，并且通过OpenCV库的resize函数进行插值放大到指定尺寸即可实现。

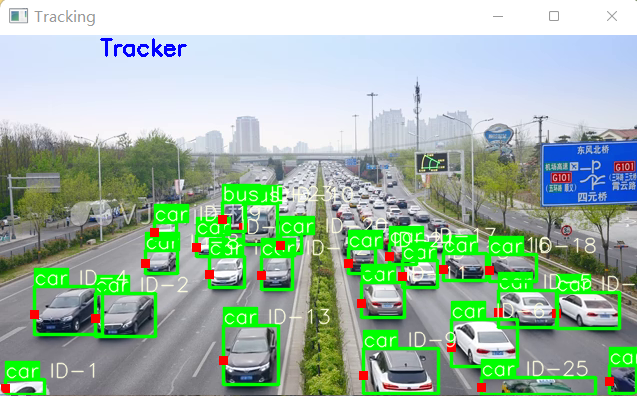
**4.6项目整体实现**

本项目运行后调用getfilepath函数获取用户所选择的文件路径，读取文件然后对视频文件的每一帧图片进行提取，对图片中的人、自行车、轿车、摩托车、公交车、卡车等物体识别与跟踪，并且对其进行编号。当用户按下‘0’键时截取该帧图片，调用selectROI函数，待用户框选目标后对目标进行跟踪同时记录下矩形框四个点的坐标，在跟踪过程中对坐标进行实时更新，在无法识别到目标后恢复到识别模式。若检测到用户按下‘2’键则对矩形框内的图像进行插值放大，检测到‘3’则退出放大模式。用户按下ESC则退出程序。

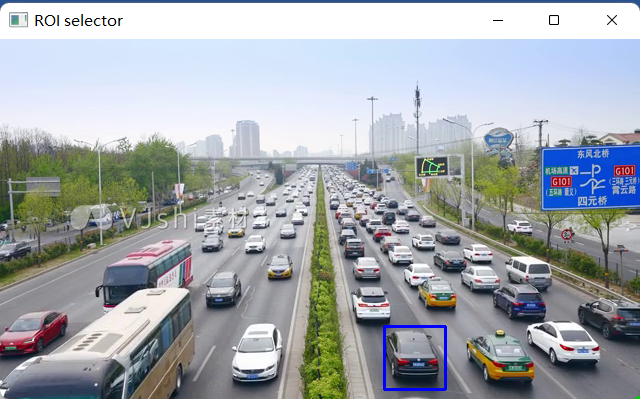
1. 项目测试
   1. 运行代码，弹出用户界面，选择视频文件。



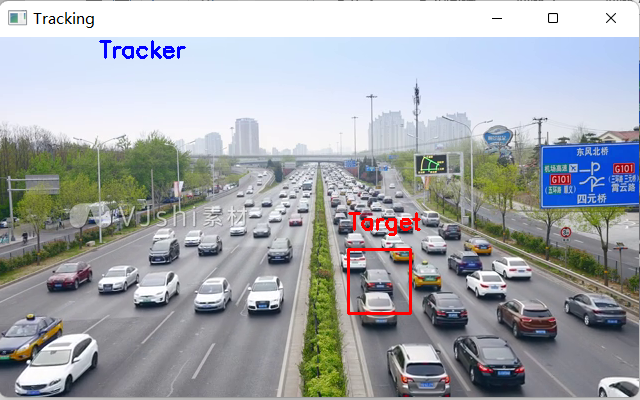
* 1. 自动进入目标识别跟踪模式，由于视频分辨率有限，导致较远处的车辆未能识别出来，在较近处的车辆识别较为准确。



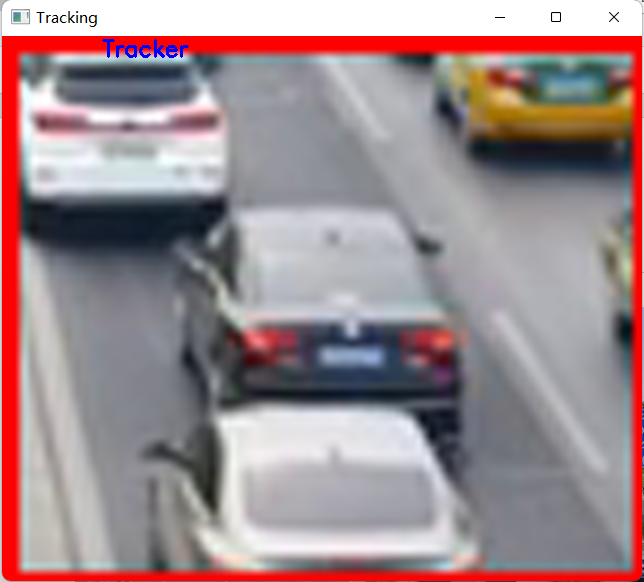
* 1. 按下‘0’，进入目标选择模式，框选如图所示的小车



* 1. 按下回车键进入单目标跟踪模式，按下‘1’可退出跟踪模式



* 1. 按下‘2’将目标放大，按‘3’恢复



* 1. 按ESC键退出程序

6 项目管理

**6.1 团队人员组成**

团队人员：陈晨宇 段一杰 鲍政儒 程璐

**6.2 任务分工**

陈晨宇 ：主要代码编写，包括main、getfilepath的编写以及tracker代码的修改，实现物体的识别、物体的跟踪、画面的展示、框选一个物体、跟踪选中的物体、放大选中的物体，贡献较大。

段一杰 ：寻找基于YOLOv5的模型并寻找资源进行测试，最终筛选出适合本项目的模型，在代码遇到错误时研究解决方案，参与报告撰写，贡献中等。

程璐 ：主要负责资料的收集和报告的撰写，将项目以书面形式展示出来，贡献较中等。

鲍政儒 ：答辩和ppt制作，将项目展示给老师同学，贡献中等。

7. 总结与反思

通过对本题目的学习，我们对于计算机视觉与计算机学习有了进一步的认识，并且更加了解了yolov5和deepsort两种算法的应用以及图像识别和追踪的原理以及具体实现，将所学的python语言应用到实际项目中是一件很有获得感的事情，这让我们更能体会到学习与应用的快乐。

本项目的优势之处在于有选择文件识别、跟踪、放大功能、图形化界面，功能齐全，且跟踪能力强，不易跟丢。本项目做的缺陷是对于较为模糊物体的识别成功率不高，同时没有生成exe文件，运行代码仍需在Python环境进行，不利于程序的推广应用。

改进的计划是尝试将代码做成一个软件，并上传到github，在更多人的使用下发现问题并不断改善。