

# 天津大学



## 《中科闻歌人工智能实训》实训报告

课题：车牌隐私保护

姓 名 何赋之 梁世豪

专 业 计算机科学与技术

指导老师 马先钦

助 教 段化芳

2021 年 9 月 26 日

## 目录

一、 概述.....	1
二、 目标检测算法发展历程.....	1
2.1 基于手工特征的传统方法.....	2
2.1.1 传统方法步骤.....	2
2.1.2 传统方法缺陷.....	2
2.2 基于深度学习的方法.....	2
2.2.1 两阶段算法.....	2
2.2.2 一阶段算法.....	5
三、 方法介绍.....	6
3.1 构建数据集.....	7
3.1.1 原始数据.....	7
3.1.2 构建本课题所用数据集.....	8
3.2 数据预处理.....	8
3.3 模型训练.....	8
3.4 车牌识别与马赛克添加.....	9
四、 实验结果.....	9
4.1 车牌识别结果.....	9
4.2 对车牌添加马赛克结果.....	10
4.3 对视频车牌进行打码结果.....	10
五、 实验总结.....	11
5.1 遇到的问题与解决方案.....	11

5.2 思考讨论.....	13
5.3 实验总结感悟.....	13
参考文献.....	13

## 一、概述

信息时代，各样的视频充斥着我们的生活，但在眼花缭乱的视频中，难免出现信息泄露问题，其中车牌信息就是值得我们注意的。本次课题任务是，对视频中的车牌进行自动识别，并将识别到的车牌添加马赛克，保护居民隐私。

要检测出车牌的准确位置，就需要使用目标检测技术。本组调研了目标检测技术发展历程，以及各类算法的优缺点，最终选择 yolov5 模型作为核心目标检测模型。

经过数据处理、增强后，得到训练好的模型。经过在后续测试中，模型性能表现良好，并针对遇到的问题实行改进。最后提出一些思考感悟。

## 二、目标检测算法发展历程

图像分类、目标检测、分割是计算机视觉领域的三大任务。在分类任务中，关注的是整张图片的内容描述，而目标检测关注特定的物体目标，要求同时获得这一目标的类别信息和位置信息。相比分类，检测给出的是对图片前景和背景的理解，需要从背景中分离出感兴趣的目标，并确定这一目标的类别和位置（常用矩形检测框的坐标表示）。

目标检测算法的发展历程，可分为基于手工特征的传统方法和基于深度学习的方法。

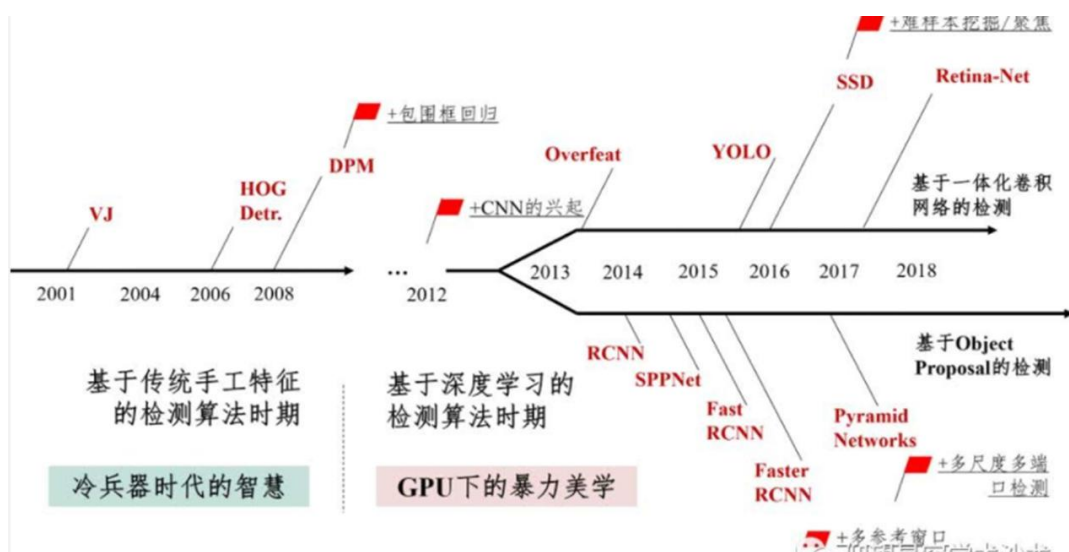


图 1 目标检测算法发展历程

## 2.1 基于手工特征的传统方法

### 2.1.1 传统方法步骤

**图像预处理：**常用的方法有高斯或均值滤波、图像腐蚀和膨胀、二值化等。

**目标区域选择：**对目标可能出现的位置进行定位。由于目标可能出现在图像中的任何位置，且目标的大小、长宽比例在一开始可能无法确定，因此最原始的方法是用不同尺寸大小的滑动窗口对全图进行遍历。

**特征提取：**对图形中目标区域的窗口进行特征提取，这个阶段是目标检测中最重要阶段，提取的特征好坏程度直接影响最后分类结果的准确率。人脸检测常用的 Harr 特征，行人检测和普通目标检测常用的 HOG 特征等。

**分类器分类：**将特征提取的结果表示成向量形式，交给特征分类器进行分类，给出所属分类的概率。常用分类器 SVM 等。

### 2.1.2 传统方法缺陷

**滑动窗口：**基于滑动窗口的区域选择策略没有针对性，时间和空间复杂度高，产生大量无用的特征。

**特征工程：**需要人工对目标区域选择合适的特征，工程时间长，经过特征工程选择的特征不一定能够符合多样性特征，其鲁棒性较差。

## 2.2 基于深度学习的方法

当 CNN 在 2012 年 ImageNet 分类任务中取得了巨大成功后，Girshick 等人便于 2014 年率先提出了区域卷积网络目标检测框架 (R-CNN)，自此目标检测领域开始以前所未有的速度发展。

目前目标检测领域的深度学习方法主要分为两类：两阶段 (Two Stages) 的目标检测算法；一阶段 (One Stage) 目标检测算法。

### 2.2.1 两阶段算法

**两阶段算法：**首先由算法生成一系列作为样本的候选框，再通过 CNN 进行样本分类。

#### R-CNN

1) 模型通过“选择性搜索”获取潜在的候选框；

- 2) 截取原图每个候选区域并 resize 输入到模型 (CNN) 中进行特征抽取;
- 3) 使用 SVM 进行分类, 以及进行 bounding box 回归

存在的问题:

- 1) 重复编码: 由于候选区域存在重叠, 模型需要重复进行重叠区域的特征图提取, 计算冗余;
- 2) 模型训练: 由于特征抽取模型和区域的分类回归模型分开训练, 无法端到端的训练, 训练过程需要提取每个包含重叠区域的候选区域特征并保存用于分类和回归训练;
- 3) 实时性差: 重复编码导致实时性不佳, 选择性搜索耗时严重。

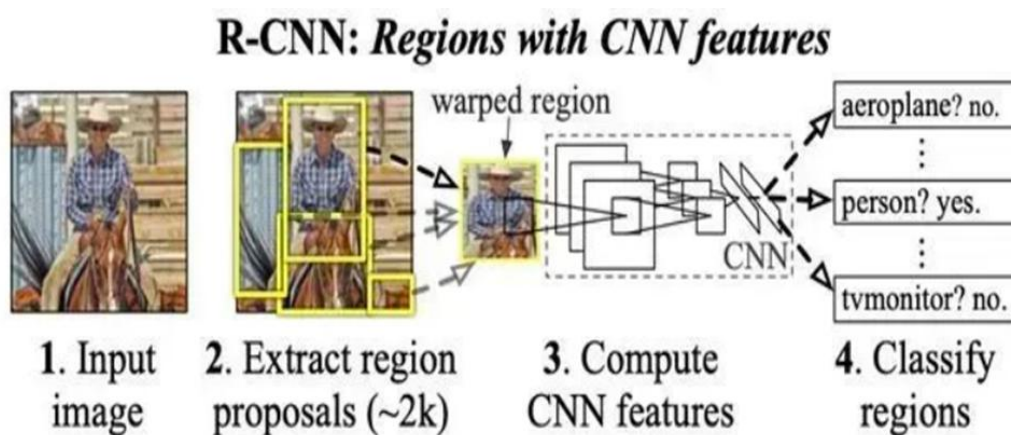


图 2 R-CNN 架构

### Fast R-CNN

- 1) 模型依旧通过“选择性搜索”获取潜在的候选框;
- 2) 将原图通过特征抽取模型进行一次的共享特征图提取, 避免了重复编码;
- 3) 在特征图中找到每一个候选框对应的区域并截取“区域特征图”, 并池化到统一大小;
- 4) 分别进行 softmax 分类和 bbox 回归。

主要优点: 可以进行端到端模型训练; 不需要存储中间特征向量用于 SVM 分类和回归模型训练。

存在问题: 选择性搜索获取候选区域耗时, 实时性差。

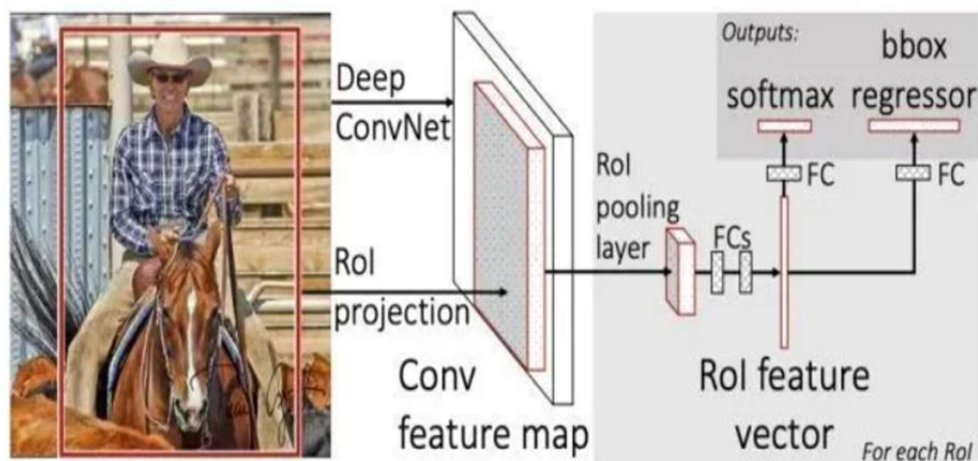


图 3 Fast-RCNN 架构

### Faster R-CNN

提出了区域生成网络（RPN），使用 RPN 网络代替选择性搜索进行候选区域的提取，相当于  $\text{Faster R-CNN} = \text{RPN} + \text{Fast RCNN}$ 。

1) 多尺度目标: 通过 RPN 网络候选区域, 并使用不同大小和长宽比的 anchors 来解决多尺度问题

2) 通过计算 anchors 与真实框的交并比 IOU, 并通过阈值建立正负样本

3) 样本不平衡: 每批次随机采样 256 个 anchors 进行边框回归训练, 并尽可能保证正负样本数相同, 避免负样本过多导致的梯度统治问题。

### Feature Pyramid Networks (FPN)

原有的目标检测算法通常都是只采用顶层特征做检测, 原因是网络顶层特征的语义信息比较丰富, 但是目标位置信息却比较粗略, 不利于目标包围框的准确定位, 相反, 虽然底层特征的语义信息比较少, 但其中目标位置信息却非常准确。

FPN 的主要思想就是在网络前馈结束后, 将网络中最顶层的特征图像逐层地反馈并与前层的特征图进行融合, 再从网络中不同深度的位置引出多个检测端口对其中的不同尺度的目标进行检测。由于网络在前馈过程中天然形成了金字塔形状的特征图, 所以 FPN 对小目标以及尺度分布较大的目标具有天然的检测优势。

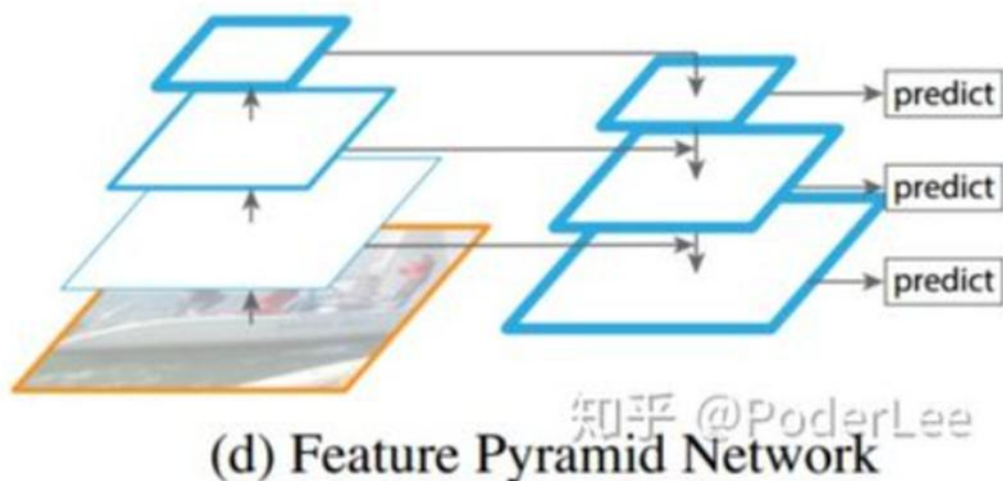


图 4 FPN 特征图融合

### 2.2.2 一阶段算法

一阶段算法：不需要产生候选框，直接将目标框定位的问题转化为回归（Regression）问题处理

#### YOLO

1) 图像划分：划分为等分的网格，物体中心落在哪个 cell，哪个 cell 就负责预测这个物体；

2) 卷积网络：每个网格上生成  $B$  个 box，每个 box 预测  $(5+C)$  回归值，四个表征位置，第五个表征这个 box 含有物体（注意不是某一类物体）的概率和位置的准确程度（由 IoU 表示）， $C$  表示类别概率；

3) 后处理：使用 NMS 过滤得到最后的预测框。

#### SSD

SSD 算法吸收了 YOLO 速度快和 RPN 定位精准的优点，采用了 RPN 中的多参考窗口技术，并进一步提出在多个分辨率的特征图上进行检测。

SSD 与 Faster-RCNN 中不同之处有两点：首先，SSD 是在多个尺度下的特征图上设置多个参考窗口进行候选框的检测，而 Faster-RCNN 仅仅是在一个尺度上进行处理；其次，Faster-RCNN 是先提取目标候选框，然后再在候选框基础上利用检测网络和 ROI Pooling 进行后续判断，SSD 则直接分别在多个尺度的特征图上进行多尺度检测和 bbox 坐标回归。



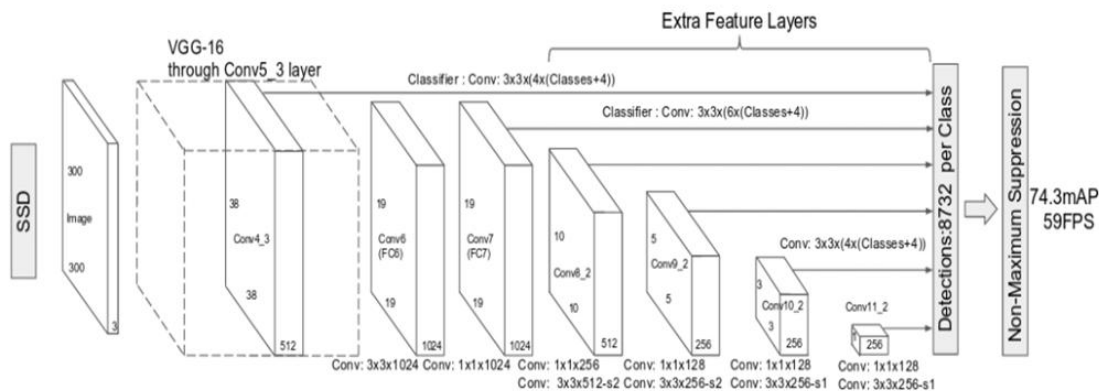


图 5 SSD 架构

### 三、方法介绍

本课题任务是，对视频中的车牌进行自动识别，并将识别到的车牌添加马赛克，保护居民隐私。经过组内讨论，认为本课题的关键在于对车牌目标的识别。因此核心在于目标检测。从模型精度、算力需求等方面考虑，本组决定使用 yolov5 作为目标检测算法。

YOLOV5 优势：

- 1) 检测精度高，适合于绝大多数数据集和目标检测任务。
- 2) 速度快，没有候选框选定环节，节省大量算力。
- 3) 使用 Pytorch 框架，对用户非常友好，能够方便地训练自己的数据集。
- 4) 不仅易于配置环境，模型训练也非常简单快速。
- 5) 代码易读，整合了大量的计算机视觉技术，非常有利于学习和借鉴。

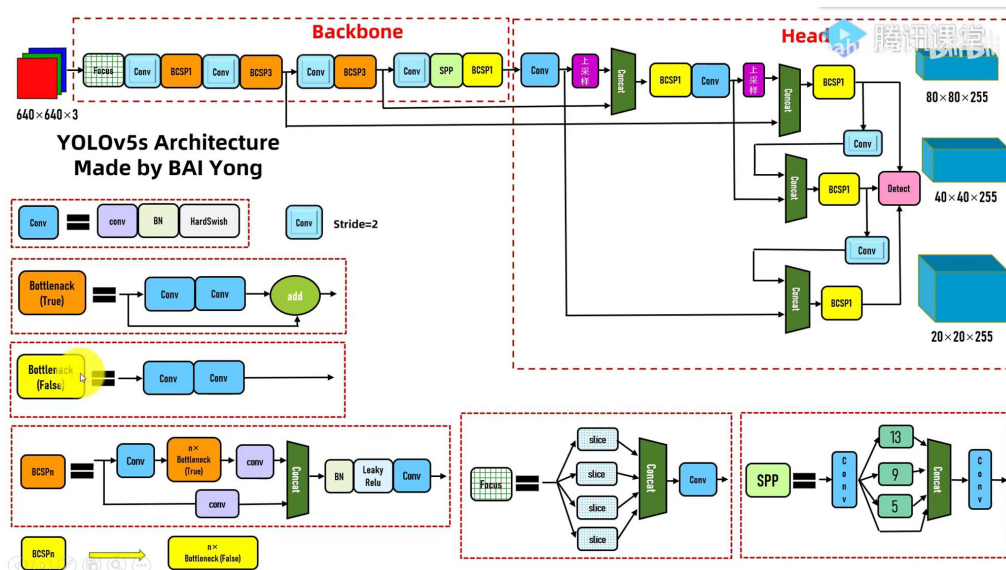


图 6 Yolov5 模型架构

方法流程图如下：

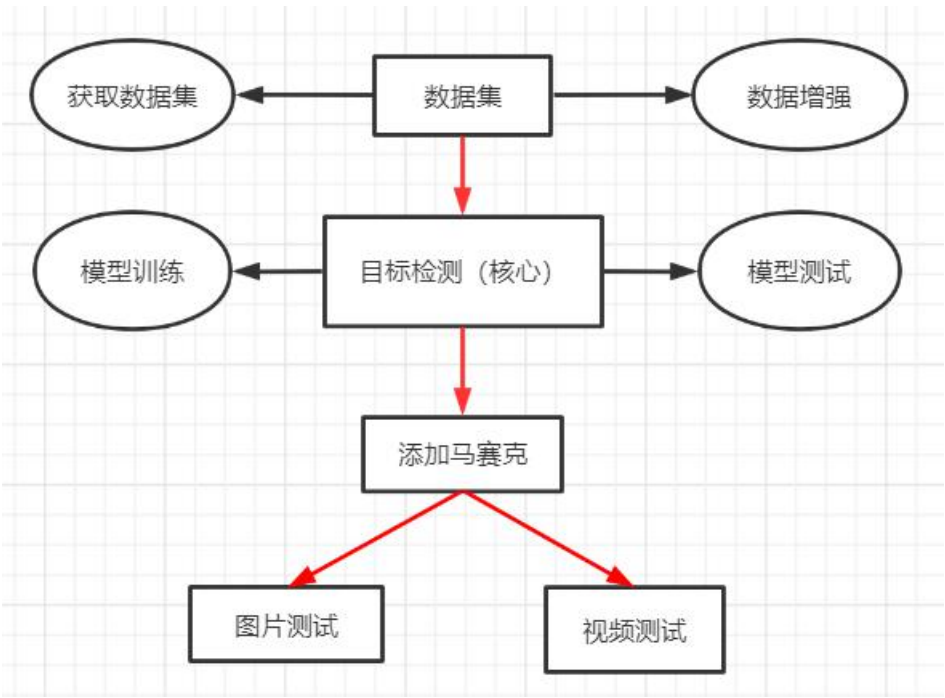


图 7 方法流程图

### 3.1 构建数据集

#### 3.1.1 原始数据

##### CCPD2019

CCPD 数据集是中科大团队建立的车牌数据集，此数据集目标用于做端到端的目标检测训练，解决了传统车牌识别检测的数据集规模较小问题。拍摄的车牌照涉及多种复杂环境，包括模糊、倾斜、阴雨天、雪天等。总数据量约 30W。

类别	描述
CCPD-Base	通用车牌图片
CCPD-FN	车牌离摄像头拍摄位置相对较近或较远
CCPD-DB	车牌区域亮度较亮、较暗或者不均匀
CCPD-Rotate	车牌水平倾斜20到50度， 竖直倾斜-10到10度
CCPD-Tilt	车牌水平倾斜15到45度， 竖直倾斜15到45度
CCPD-Weather	车牌在雨雪雾天气拍摄得到
CCPD-Challenge	在车牌检测识别任务中较有挑战性的图片
CCPD-Blur	由于摄像机镜头抖动导致的模糊车牌图片

图 8 CCPD2019 各种类数据

## bluegreenyello

CCPD2019 数据集仅有蓝色车牌，于是从网上获取该数据集，从而加入部分绿色车牌（204 张）与黄色车牌（650 张），扩充了蓝色车牌（882 张）

### 3.1.2 构建本课题所用数据集

从 CCPD2019 每种图片类型中抽取十分之一，获取约数据量 3W 的数据，并加入所有 bluegreenyello 数据集中的数据。同时注意到，尽管 CCPD2019 中有对车牌进行旋转处理，但旋转角度较小，以此训练的模型在遇到特殊情况下的竖直车牌时，可能表现不佳，于是对部分车牌进行大幅度旋转操作（约 5000 张），进行数据增强。加入旋转图片中，组成最终的项目所用数据集，共计约 4W 张图片。

## 3.2 数据预处理

本课题所用的 yolov5 模型，需要 voc 类型数据，但在 CCPD2019 中仅含图片，且标注信息蕴含在图片名称中，需要进行进一步处理，生成需要的 xml 文件。具体代码处理过程详见 readme.md 文件中。

## 3.3 模型训练

模型参数：Yolov5 子架构：yolov5-l；epoch：200；img-size：[320, 320]；batchsize：32。

模型性能：

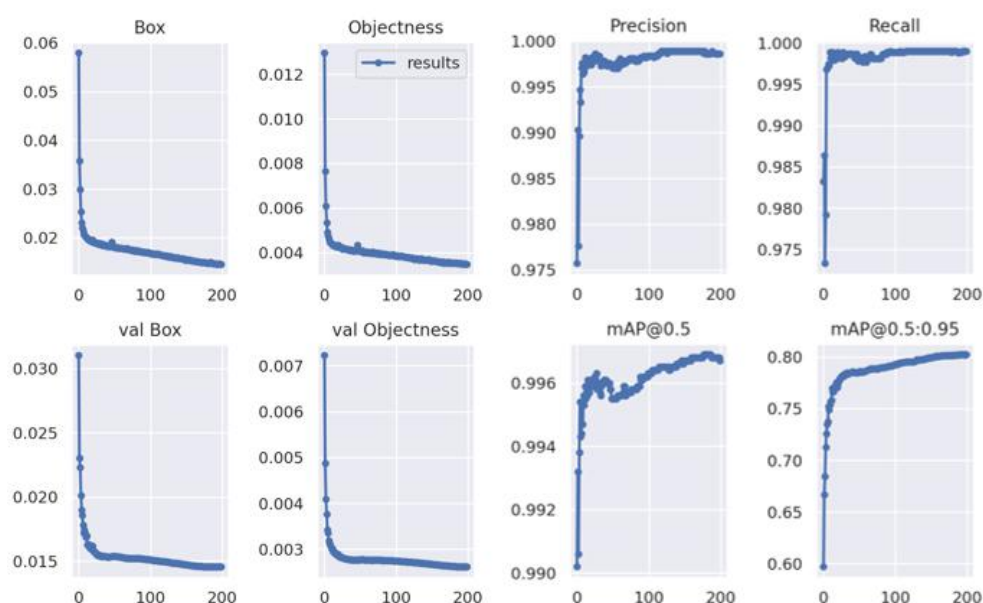


图 9 训练出的 yolov5 模型性能

### 3.4 车牌识别与马赛克添加

加载训练得到的表现最好的权重参数 `best.pt`，用以在图片/视频上进行检测。识别出来后，模型返回预测框的左上角 `x`、`y` 坐标与右下角的 `x`、`y` 坐标，调用 `cv2.rectangle` 函数，将预测框中的 RGB 像素值改为 `(0, 0, 0)`，则可达到马赛克的效果，最终效果为一个全黑的方框。将函数封装成方法供主检测程序 `detect.py` 使用，即可在原图上画出预测框并对其打码。

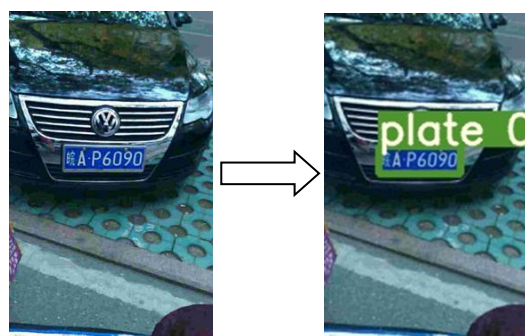
另外，即使训练数据中没有视频，但对于视频的检测时可根据视频长度将其分为若干帧，即可达到预期检测效果。

## 四、实验结果

### 4.1 车牌识别结果

由图 9 可知，训练的 `yolov5` 模型在精确率、召回率等各项指标上，都表现出了强大的性能。该模型在后续测试中的车牌检测结果也十分优异，无论在 `bbox` 位置还是置信度方面均表现较好。具体实例如下：

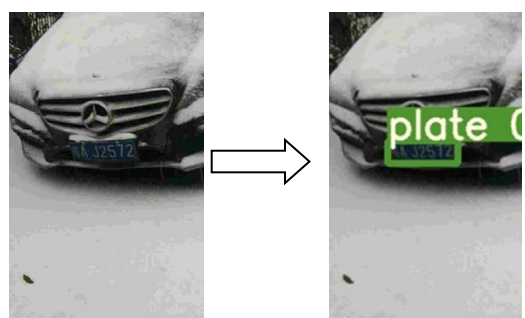
常规环境：



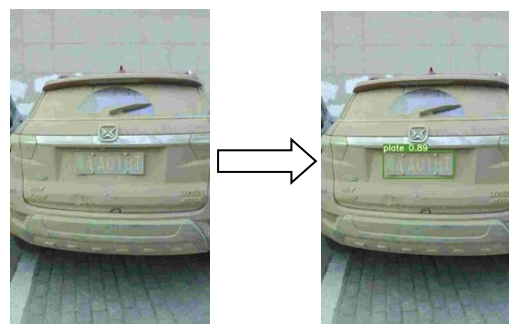
黑夜：



特殊天气：



车牌模糊、磨损：



其它颜色车牌，多目标车牌：



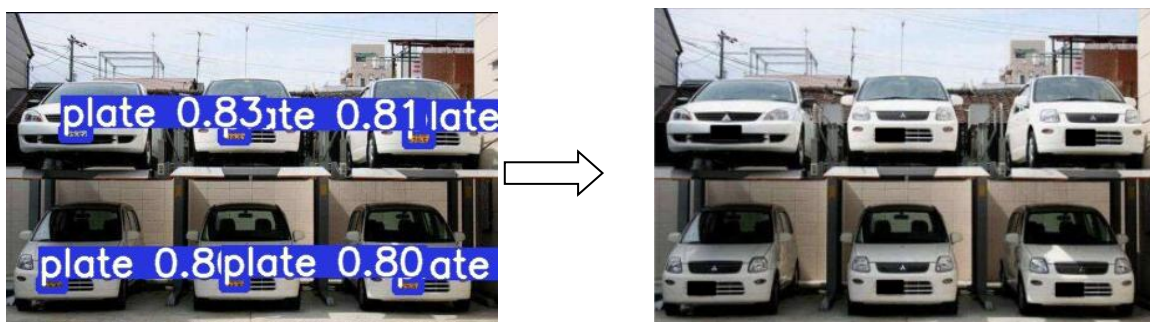


带有旋转的车牌：



## 4.2 对车牌添加马赛克结果

利用 yolov5 模型准确获取到车牌的位置后，即可进行打码操作。在与老师沟通交流后，为避免对部分图片添加常规马赛克时遇到的车牌像素过高，添加马赛克后仍可清晰分辨车牌号的问题，本组决定将打码操作简单化，直接对检测区域实行遮盖处理，实现隐私保护任务。



## 4.3 对视频车牌进行打码结果

大致效果如下图：



可见，对于近景和中景下的复杂场景，模型都能较好地检测出车牌。  
完整结果见压缩包。

## 五、实验总结

### 5.1 遇到的问题与解决方案

#### 1) 对非车牌的错检

在测试过程中发现，存在少量的错检现象，模型有一定概率将图片中的网站、电视台名称错检为车牌。

解决方案：发现这类错检的置信度较低，提高模型检测置信度阈值即可。



#### 2) 对竖直车牌的漏检

在未加入大幅旋转的车牌图片时，对某些特殊情况下的竖直车牌产生漏检。

解决方案：遵循老师的建议，对部分图片进行旋转操作，增加旋转图片，增强数据集。



### 3) 对小目标车牌的漏检

在将模型运用至视频检测后发现,对视频中的较远距离的小车牌容易发生漏检现象。

解决方案:训练模型时,增大 image-size,使模型训练时更容易提取小目标特征,在后续视频测试中该问题得到较好改善。

## 5.2 思考讨论

(1) 训练数据的质量在整个工程中十分重要,直接影响模型训练后的效果。通过解决竖直类车牌漏检问题,学习到了数据集需要涵盖实际情况的方方面面,充分考虑各类可能遇到的情况,这样训练出的模型才具有更好的鲁棒性。

(2) YOLOv5 模型的表现令人满意。对于照片、视频都能较好地检测出车牌,作为轻量级模型,训练时间短,效果好,可以解决实际生活中大部分目标检测问题。

(3) 在小目标检测上仍可以进行改进。训练集中小目标车牌样本十分稀少,可以手工添加小目标车牌,达到更好的训练效果。

## 5.3 实验总结感悟

经过为期一个多月的实训,本组成员体会到了课题项目和平时学校课题作业的不同。课题项目需要考虑诸多实际情况,在整个了解任务-制定方案-实施方案-解决问题的过程中,我们得到了很好的锻炼,大大提高了解决实际问题的能力,深刻体会到理论联系实际的重要性。感谢中科闻歌提供本次宝贵的实训机会,本组成员定认真总结本次实训的收获,让其成为一笔高贵财富,也祝中科闻歌一帆风顺,越办越好。

## 参考文献

- [1] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [2] Alexey Bochkovskiy, Chien-Yao Wang, et al. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [3] Zhengxia Zou, Zhenwei Shi, Yuhong Guo, Jieping Ye. Object Detection in 20 Years: A Survey. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019.
- [4] Hui Li, Peng Wang, Chunhua Shen et al. Towards End-to-End Car License Plates Detection and Recognition with Deep Neural Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
- [5] Z Xu, Y Wei, A Meng, N Lu, L Huang et al. Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline. In *European Conference on Computer Vision(ECCV)*. 2018.