

基于YOLOv5的车牌识别系统

崔兆文, 刘肖飞*, 夏诗楠, 瞿 竟

华东光电集成器件研究所, 江苏 苏州

收稿日期: 2023年6月9日; 录用日期: 2023年7月18日; 发布日期: 2023年7月27日

摘 要

近年来, 车辆的数量不断增多, 智能交通场景的不断完善, 在现实生活中车牌识别技术的重要性显而易见, 本文提出了一种YOLOv5网络与LPRNet网络相结合的技术, 同时使用透视转化的方法对车牌进行校正, 进而实现对车牌的定位与识别, 该模型的模型召回率在99.7%, 准确度84.4%, 能够准确识别车牌号。本文先对研究的背景和内容进行了简单的介绍, 接着对相关理论内容进行了概述, 然后分析实验数据与结果, 最后对本次研究进行总结。

关键词

车牌识别系统, YOLOv5, LPRNet, 透视转化

A License Plate Recognition System Based on YOLOv5

Zhaowen Cui, Xiaofei Liu*, Shinan Xia, Jing Qu

East China Institute of Optoelectronic Integrated Devices, Suzhou Jiangsu

Received: Jun. 9th, 2023; accepted: Jul. 18th, 2023; published: Jul. 27th, 2023

Abstract

In recent years, the number of vehicles has not increased significantly, and the continuous improvement of intelligent transportation scenarios has made the importance of license plate recognition technology obvious in real life. This article proposes a technology that combines YOLOv5 network and LPRNet network, and uses perspective transformation method to correct the license plate, thereby achieving the positioning and recognition of the license plate. The model has a recall rate of 99.7%, an accuracy of 84.4%, and can accurately recognize the license plate. This ar-

*通讯作者。

title first provides a brief introduction to the background and content of the research, followed by an overview of relevant theoretical content, analysis of experimental data and results, and finally a summary of this study.

Keywords

License Plate Recognition System, YOLOv5, LPRNet, Perspective Transformation

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

发展智慧城市是中国促进城市高度信息化、网络化的重大举措。其中，通过监控对行驶中的车辆进行实时的车牌信息检测是智慧交通中的一个重要分支。

在机动车牌识别方面国内外的研究者们做了相关的研究。Saini 等通过一种多小波变换进行定位车牌的框架，采用了图像增强以及畸变校正，提高了框架的有效性[1]，但是这种传统的方法受环境干扰较大。顾李云通过高斯与中值滤波进行图像预处理，提高了传统算法在弱光条件下处理图像的能力，与传统的 Sobel 算子相比，通过 Mallat 小波变换提取图像边缘更加完整[2]，但是对于车牌图像的质量以及车牌的位置要求较高，对于质量差以及未校正的车牌识别较差。王宁基于 Faster R-CNN (Faster Region-based Convolutional Neural Network)在 RPN 阶段引入角度和分支结构以生成适合车牌的候选区域，并提出了一种上下文融合网络增强车牌的特征表示[3]，Faster R-CNN 的模型较大，在同样环境下识别时间较长。周世杰基于卷积神经网络提出了一种 textboxes++网络结构进行车牌检测[4]，该方法对车牌定位效果较差，对图片质量要求较高。陈冬英等人研究改进 LeNet5-CNN (LeNet-5 Convolutional Neural Network)使其在车牌的识别率上有了了一定的提高[5]，但是卷积神经网络的运行速度还是影响了网络运行的速度。

为了解决上述研究者存在的问题，本文提出一种基于 YOLOv5 (You Only Look Once v5)与 LPRNet (License Plate Recognition Network)技术相结合方式实现车牌的识别，使用 YOLOv5 定位车牌的位置，同时使用 LPRNet 技术实现对车牌号的识别。该方法通过训练的方法对各类车牌进行定位，尽可能地减少周围环境对车牌定位的影响，使用透视变换的方法对图片进行校正，使得 LPRNet 技术识别出来的车牌更加准确。

2. 理论与方法

深度学习在图片识别的应用领域越来越广泛，本文基于深度学习的方法进行机动车车牌相关算法的研究，为了提高算法的运行效率，本文选取一次性预测的方法 YOLO，作为车牌定位的基本方法，确保车牌定位不出现较大偏差，然后通过透视变换的方法对车牌进行相关校正，将图片中倾斜的车牌尽可能校正还原，最后通过 LPRNet 技术实现对这牌信息的读取。

2.1. YOLOv5 算法

YOLOv5 是由 Glenn Jocher 等基于 YOLO 优化改进而来，模型分为四个部分，分别是输入端、主干网络、颈部网络以及预测网络组成，其模型框图如图 1 所示。

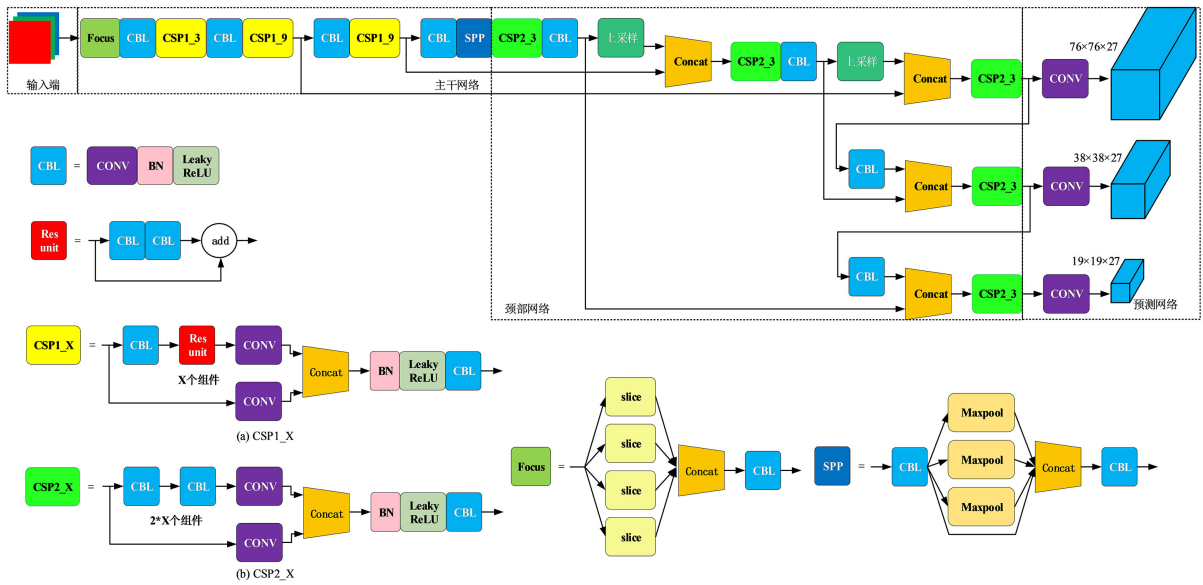


Figure 1. YOLOv5 Model Block Diagram
图 1. YOLOv5 模型框图

1) 输入端

输入端包括数据增强、自适应锚框计算、图片尺寸处理等。

数据增强采用 Mosaic 数据增强的方式，即随机缩放、随机裁剪、随机排布的方式进行拼接。通过这种数据增强的方式，增加很多小目标，提升小物体的识别能力，符合本文机动车车牌这种小目标检测需求。

在 YOLO 算法中，网络训练前都需要设定初始的锚框，网络在训练时候，根据初始锚框来进行输出预测框，然后通过与真实框相比对，根据两者计算的差值进行反向更新，进而反复迭代更新网络参数。

在 YOLO 系列算法中，输入图像的需要统一尺寸大小，然后输入检测模型中进行训练，YOLOv5 在输入端提供了图像缩放的转化方法，输入模型的图像可以不转化为固定的大小，避免因为填充方法的不同造成信息冗余等不良影响。

2) 主干网络

主干网络由 Focus 结构以及 CSP (Cross Stage Partial Network)结构构成。假设把原始图像 $608 \times 608 \times 3$ 经由切片操作，输出 $304 \times 304 \times 12$ 大小的特征图，随后通过 32 个卷积核进行一次卷积操作，变为 $304 \times 304 \times 32$ 的特征图。

YOLOv5 中包含了两类 CSP 结构，分别为 CSP1_X 以及 CSP2_X。在网络模型中，主干网络主要使用了 CSP1_X 结构，而 Neck 结构中主要使用了 CSP2_X 结构。其中 CSP1_X 模块由残差神经网络与 CBL (Conv Batch normalization Leaky relu)模块组成，CSP2_X 由两个 CBL 模块组成，CSP1_X 模块主要应用在主干网络中用于加强特征提取，CSP1_X 主要应用在颈部网络中，用于加强特征融合。具体结构在图 1 中有体现。

3) 颈部网络

FPN + PAN 的网络结构被应用于预测层中，从上到下的 FPN 结构通过上采样以及从上到下的 PAN 下采样来处理信息传递过程，以得到预测的特征图。从下到上的特征金字塔应用在 PAN 结构中。具体形式如图 2 所示。

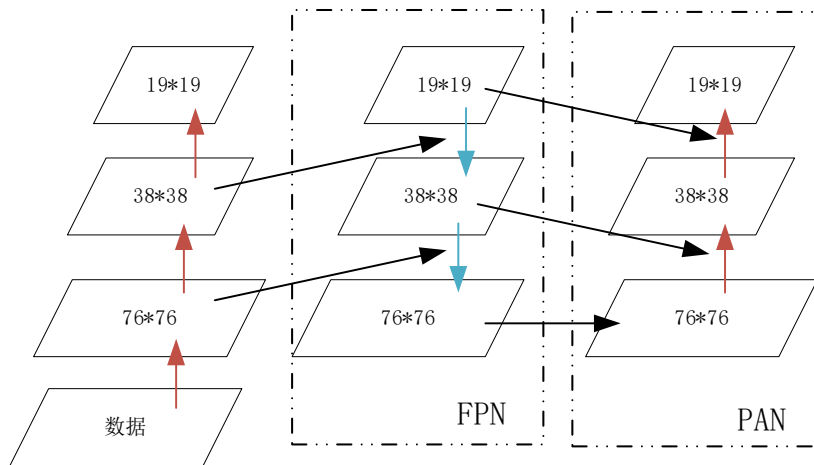


Figure 2. The Network Structure of FPN + PAN
图 2. FPN+PAN 的网络结构

4) 预测网络

YOLOv5 的损失函数由分类损失、定位损失、置信度损失三部分组成，损失函数计算公式如式(1)所示，其中 $Loss_{cla}$ 表示分类损失部分， $Loss_{loc}$ 表示定位损失部分， $Loss_{con}$ 表示置信度损失部分。其中定位损失函数方面使用的是 $GIoU$ (Generalized Intersection over Union)方式实现的。防止出现以下两种情况：

- 1) 当真实框和预测框完全不重叠时，交并比损失值恒定，无法完成损失值的更新；
- 2) 当真实框包含预测框，交并比损失值恒定，无法完成损失值的更新。

$$Loss = Loss_{cla} + Loss_{loc} + Loss_{con} \tag{1}$$

同时使用非极大值抑制 NMS (Non-Maximum Suppression)，确保一张车牌不会反复地被识别为车牌，一个车牌只识别一次。

2.2. 透视变换

透视变换是将图片投影到一个新的视平面，也称作投影映射。通用的变换公式为式(2)。

$$[x', y', z'] = [u, v, w] \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{2}$$

u, v 是原始图片左边，对应得到变换后的图片坐标 x, y ，其中 $x = x'/w'$ ， $y = y'/w'$ 。

变换矩阵 $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ 可以拆分为四个部分， $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 表示线性变换 $[a_{31} \ a_{32}]$ 用于平移，

$[a_{13} \ a_{23}]^T$ 产生透视变换。可以得到式(3)，一般 $a_{33} = 1$ 则方程中有 8 个未知数，找 4 个点即可完成参数的求解。

$$\begin{aligned} x &= \frac{a_{11}u + a_{21}v + a_{31}}{a_{13}u + a_{23}v + a_{33}} \\ y &= \frac{a_{12}u + a_{22}v + a_{32}}{a_{13}u + a_{23}v + a_{33}} \end{aligned} \tag{3}$$

2.3. LPRNet

LPRNet 网络是卷积神经网络轻量化的模型, 无需字符分割便实现端到端的车牌字符识别。网络结构设计包括: 具有空间变换器层的定位网络、轻量卷积神经网络、每个位置的字符分类头、序列解码的字符概率、后过滤过程等。其主干网络结构如表 1 所示, CNN (Convolutional Neural Network)模型主干的小基本构建块如表 2 所示。使用 CTC 损失函数(Connectionist Temporal Classification)作为 LPRNet 网络的损失函数。

Table 1. Backbone network structure

表 1. 主干网络结构

网络层	参数
输入层	94 × 24 像素 RGB 图像
卷积层	#64 3 × 3 步长为 1
最大池化层	#64 3 × 3 步长为 1
小基本构建块	#128 3 × 3 步长为 1
最大池化层	#64 3 × 3 步长为(2, 1)
小基本构建块	#256 3 × 3 步长为 1
小基本构建块	#256 3 × 3 步长为 1
最大池化层	#64 3 × 3 步长为(2, 1)
Dropout	0.5
卷积层	#256 4 × 1 步长为 1
Dropout	0.5
卷积层	# 1 × 13 步长为 1

Table 2. Small basic building blocks

表 2. 小基本构建块

网络层	参数
输入层	$C_{in} \times H \times W$ 特征图
卷积层	$C_{out}/4 \ 1 \times 1$ 步长为 1
卷积层	#64 3 × 3 步长为 1
卷积层	#128 3 × 3 步长为 1
卷积层	#64 3 × 3 步长为(2, 1)
输出层	#256 3 × 3 步长为 1

3. 实验设计与结果分析

3.1. 实验环境

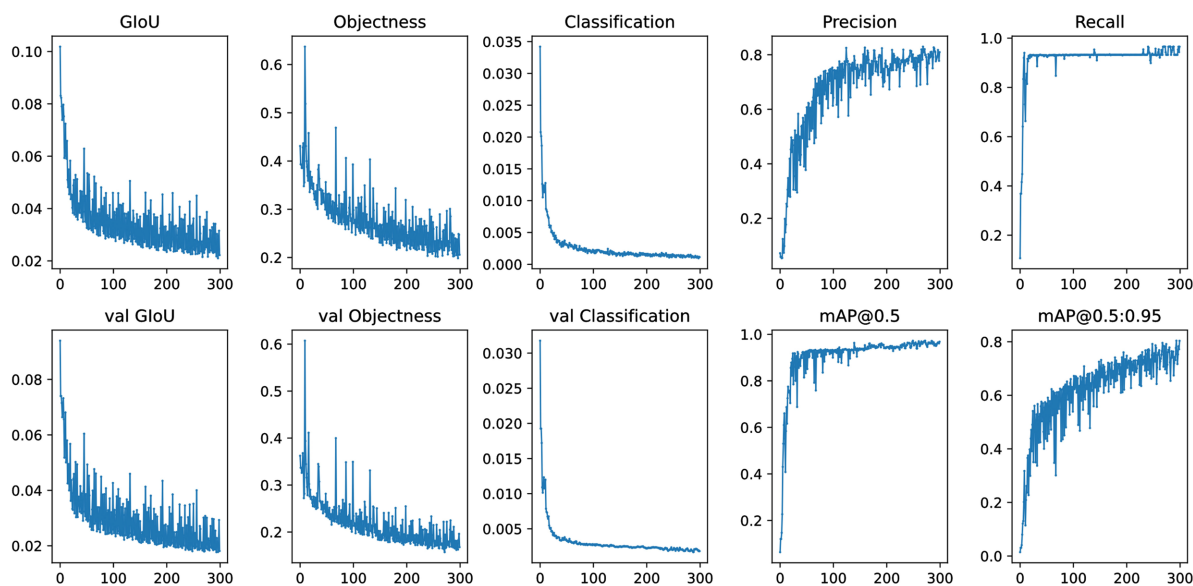
车牌识别环境如表 3 所示。

Table 3. Experimental environmental parameters**表 3.** 实验环境参数

项目	参数
操作系统	Windows 11
CPU	i5-12500H
内存	16G
系统类型	64 位操作系统

3.2. YOLOv5 训练结果

实验的机动车车牌数据来源于网络。如图 3 所示, 为 YOLOv5 训练的结果图, Glou、Objectness、Classification 为三个损失函数, 分别对应定位损失函数、置信度损失函数、分类损失函数, 其中 val 表示测试集, 没有 val 表示数据集。Recall 表示检测准确的结果在所有的正样本中的占有比例, Precision 表示正确的结果所占模型检测出的结果中的比例。mAP@0.5 表示 IoU (Intersection over Union) 阈值大于 0.5 的平均检测精度, mAP@0.5:0.95 表示 IoU 阈值在 0.5~0.95 之间等间距取 10 组的平均检测精度再取平均值。

**Figure 3.** YOLOv5 training result graph**图 3.** YOLOv5 训练结果图

本文对数据进行了 300 次的训练, 得到的 YOLOv5 车牌定位模型的召回率在 99.7%, 准确度 84.4%。虽然准确度稍低, 但是召回率很高。与 Faster R-CNN 模型的实验数据对比结果如表 4 所示, 虽然准确率上 Faster R-CNN 高于 YOLOv5, 但是无论在召回率还是算法运行速度方面 YOLOv5 优于 Faster R-CNN 更多。

Table 4. Comparison table of different models**表 4.** 不同模型对比表

检测算法	召回率	准确率	FPS
YOLOv5	99.7	84.4	42
Faster R-CNN	84.2	91.6	15

3.3. 车牌识别效果

本文对具有一定倾斜的车牌进行识别，图 4 对识别效果进行了展示，分别对蓝牌车、黄牌车以及绿牌车进行了识别，均能正确识别车牌信息。



Figure 4. Diagram of the practical teaching system of automation major
图 4. 车牌识别结果

通过本文建立的模型不仅可以显示出车牌的信息，对车牌的位置置信度也有显示。

4. 结束语

本文对机动车的车牌进行识别，对一定内倾斜的车牌有一定的辨识效果，当内倾斜角度较大时，容易造成省份名称识别不够准确，容易被识别成字母等信息，同时对于货车等两行车牌信息的识别效果较差，并未在图 4 中体现，后期可以对算法进行完善，能够识别更多的车辆。同时后期可以增加样本空间与算法迭代实现对电动车等其他车辆车牌的识别。

本次车牌识别并未建立上位机界面，仅仅只是对算法的验证，后期可以加入操作界面，或引入摄像头，对车牌信息进行实施的监控，完善汽车车牌识别系统。

参考文献

- [1] Saini, M.K. and Saini, S. (2017) Multiwavelet Transform Based License Plate Detection. *Journal of Visual Communication and Image Representation*, **44**, 128-138. <https://doi.org/10.1016/j.jvcir.2017.01.003>
- [2] 顾李云. 基于图像处理的车牌识别算法的研究与设计[D]: [硕士学位论文]. 南京: 南京邮电大学, 2018.
- [3] 王宁. 非受限场景下的车牌识别系统的研究与实现[D]: [硕士学位论文]. 南京: 南京邮电大学, 2020.
- [4] 周世杰, 李顶根. 基于卷积神经网络的大场景下车牌识别[J]. *计算机工程与设计*, 2020, 41(9): 2592-2596.
- [5] 陈冬英, 黄悦, 李凌懿, 王杰, 高泽晖, 舒萍. 改进 LeNet5-CNN 的车牌识别工程应用优化设计[J]. *武夷学院学报*, 2022, 41(12), 33-38.