

基于深度学习的车距识别研究

崔兆文, 王旭亚, 刘肖飞*, 夏诗楠

华东光电集成器件研究所, 江苏 苏州

收稿日期: 2025年7月4日; 录用日期: 2025年8月9日; 发布日期: 2025年8月18日

摘要

本文提出了一种基于深度学习的车距检测的研究, 利用YOLOv5实现对车辆的定位, 利用距离检测方式实现对车辆距离的计算, 主要介绍了YOLOv5与实现对车辆距离的计算相关理论与实验结果, 实验结果车辆判定的准确率95.5%, 测距识别精度为10%以内, 该研究对辅助驾驶技术有一定的借鉴意义。

关键词

手势识别, YOLOv5, 距离检测, 测距识别

Deep Learning-Based Vehicle Distance Recognition Research

Zhaowen Cui, Xuya Wang, Xiaofei Liu*, Shinan Xia

East China Institute of Optoelectronic Integrated Devices, Suzhou Jiangsu

Received: Jul. 4th, 2025; accepted: Aug. 9th, 2025; published: Aug. 18th, 2025

Abstract

This paper presents a deep learning-based vehicle distance detection study, which utilizes YOLOv5 for vehicle localization and implements distance measurement through geometric calculation methods. The research primarily introduces the theoretical framework of YOLOv5 and its application in vehicle distance computation, along with experimental results. The experiments achieved 95.5% accuracy in ranging identification and 10% precision in distance measurement. This study provides valuable references for advanced driver-assistance systems (ADAS) technology.

Keywords

License Plate Gesture Recognition, YOLOv5, Distance Detection, Ranging Identification

*通讯作者。

Copyright © 2025 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

随着城市化进程加速和机动车保有量的快速增长,交通拥堵、交通事故频发等问题日益严重。为了解决这些问题,智能交通系统(ITS)逐渐成为研究热点。在ITS中,车辆识别技术是核心组成部分,它能够实现交通流量监控、违章行为检测、自动驾驶辅助等功能。近年来,深度学习技术在图像分类和目标检测任务中取得了突破性进展,极大地推动了车辆识别技术的发展。

传统的车辆识别方法主要依赖于手工特征提取和浅层分类器。例如 Haar 特征、HOG 特征、SIFT 特征等被广泛应用于车辆检测任务中,结合 SVM、Adaboost 等分类算法进行识别。然而,这些方法存在较高的局限性对光照变化敏感:在不同光照条件下识别准确率波动较大;难以应对遮挡问题:面对部分遮挡时鲁棒性较差;计算效率低:处理高分辨率图像时速度较慢;泛化能力不足:在复杂背景或天气突变情况下性能下降明显。

孙伟业等人提出一种增广补零的改进 KM 算法,实现车辆识别与跟踪技术[1]。骆绍焯等人提出基于深度学习的车辆特征识别系统,能够迅速掌握车辆的各种特征,为车辆的智能化管理提供支持[2]。

本文使用 YOLOv5 (You Only Look Once version 5)与距离估测算法实现对于车辆距离的识别。

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手部特征信息提取的研究[3],为了提高算法的运行效率,本文选取一次性预测的方法 YOLO,作为车辆信息提取的基本方法,确保对车辆的定位,通过距离估测算法实现对车辆距离的判定。

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[4]。

1) 输入端。

Mosaic 数据增强是 YOLOv5 输入端最核心的技术创新,通过将 4 张训练图像随机缩放、裁剪和排布组合成一张图像:大幅提升小目标检测能力(最多可包含 4 倍于原始图像的小目标);增强模型对目标不完整情况的识别鲁棒性;实现更高效的正负样本均衡;减少 GPU 显存需求(批量归一化时能同时看到 4 张图像的内容)。

YOLOv5 在输入端实现了锚框(anchor)的自动优化:训练前自动分析数据集标注框的宽高分布;使用 k-means 聚类算法计算最优锚框尺寸;根据模型不同深度版本(s/m/l/x)调整锚框数量;支持自定义数据集的特异性锚框优化。

传统方法采用固定尺寸缩放会导致信息丢失或计算冗余,YOLOv5 改进为:保持原始图像长宽比进行缩放;自动计算最小填充(padding)量;实现“最小黑边”缩放策略;显著减少计算量(最高可减少 40% 的无效计算)。

自动调整对比度/亮度:随机调整系数在 0.5~1.5 之间,随机水平翻转:概率设置为 0.5 [5]。

2) 主干网络。

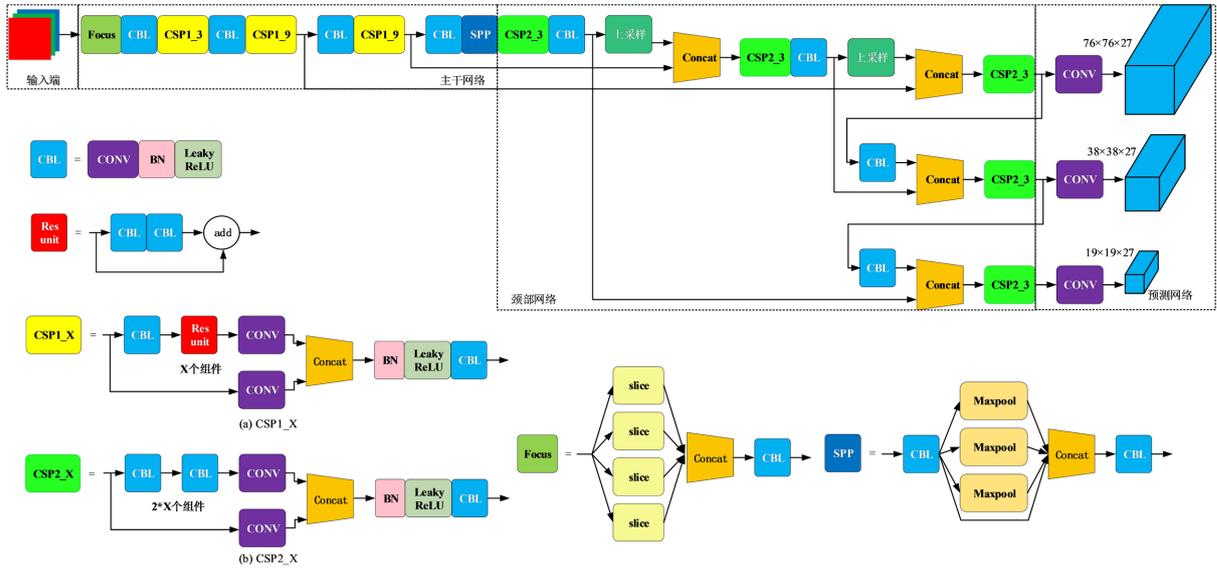


Figure 1. YOLOv5 model block diagram
图 1. YOLOv5 模型框图

YOLOv5 的主干网络是其目标检测架构的核心特征提取器，采用精心设计的 CNN 结构实现高效的多尺度特征提取。YOLOv5 主干网络主要由以下部分组成：Focus 模块(独特的切片卷积结构)、CSPNet 结构(跨阶段部分网络设计)、SPP 模块(空间金字塔池化)、Neck 连接(与特征金字塔网络(FPN)的接口)。

Focus 模块特点：通过切片操作将空间信息转换为通道信息、无信息丢失的下采样方式、计算效率比常规卷积更高。CSPNet 结构特点：由 3 个卷积组成的残差块、部分特征直接跨阶段传递、深度可分离卷积：减少计算量、梯度分流缓解梯度消失问题。

SPP 模块特点：即空间金字塔池化，多尺度特征融合能力、增大感受野而不丢失分辨率、对目标尺度变化更鲁棒、保持固定长度输出。

激活函数改进使用 SiLU(Swish)激活函数替代 LeakyReLU，其计算公式为 $SiLU(x) = x * sigmoid(x)$ [6]。

3) 颈部网络。

YOLOv5 的颈部网络主要负责特征融合，其架构设计融合了 FPN (Feature Pyramid Network)和 PAN (Path Aggregation Network)两种经典结构的优点，形成高效的 FPN + PAN 结构。具体形式如下图 2 所示。

FPN 部分：自顶向下路径，深层语义信息传递到浅层；横向连接，通过 1×1 卷积调整特征图通道数；特征融合：采用逐元素相加方式。

PAN 部分：自底向上路径，将浅层位置信息传递到深层；特征优化，使用 CSP (Cross Stage Partial Network)结构加强特征融合能力；跨尺度连接，构建完整的信息传播路径。

YOLOv5 将原始 PAN 中的卷积模块替换为 CSP 结构：采用 CSP2_1 结构，将基础特征分为两部分进行处理，最后合并两部分特征，显著减少计算量的同时保持性能。

设计了 3 层特征金字塔：P3 (80×80)：高分辨率，检测小目标；P4 (40×40)：中等分辨率，检测中等目标；P5 (20×20)：低分辨率，检测大目标。

颈部网络的作用是整合来自不同层的特征，增强模型对不同尺度和复杂背景下的目标检测能力。这些结构通常会提升模型的性能，特别是在处理多尺度目标和复杂场景时[7]。

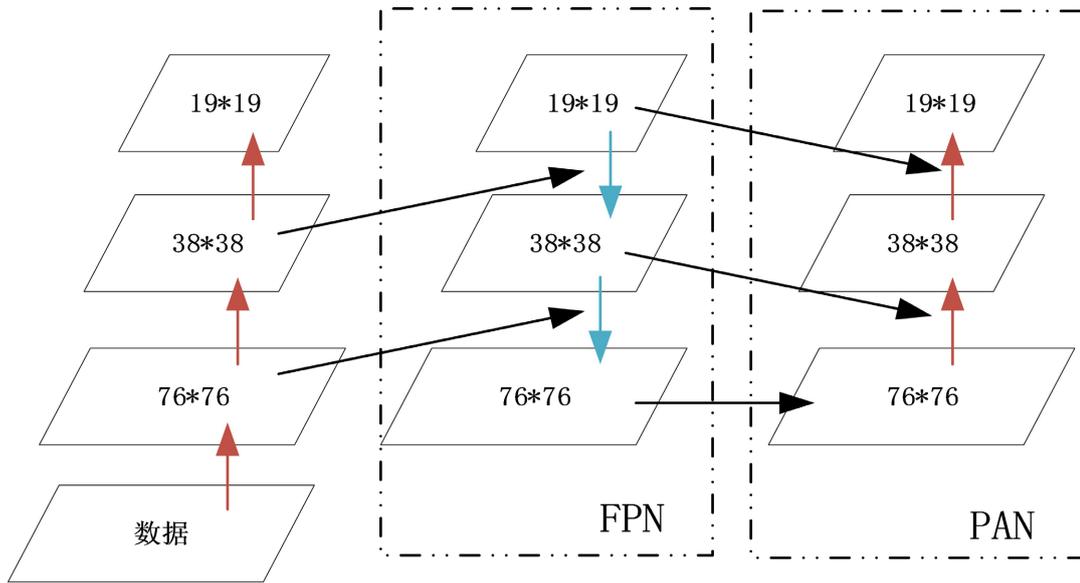


Figure 2. The network structure of FPN + PAN
图 2. FPN + PAN 的网络结构

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， λ_1 、 λ_2 、 λ_3 是损失权重，用于平衡不同类型的损失对总损失的影响。分类损失对于每个锚框(anchor)，模型预测一个类别概率分布，分类损失衡量的是预测类别与真实类别之间的差异。通常使用二元交叉熵损失函数计算，对于多类别问题，可以使用多类别交叉熵损失。定位损失衡量预测边界框的坐标与真实边界框之间的差距。YOLOv5 使用的是 GIoU 损失，这是一种改进的 IoU 损失，即使在框不重叠时也能提供梯度，有助于模型更好地学习目标位置。模型为每个锚框预测一个置信度分数，表示该框是否包含一个对象。置信度损失也是使用二元交叉熵损失计算的，它区分了预测框是否包含目标物体。

$$Loss = \lambda_1 Loss_{cla} + \lambda_2 Loss_{loc} + \lambda_3 Loss_{con} \quad (1)$$

同时使用非极大值抑制 NMS (Non-Maximum Suppression)，确保手的位置不会被反复识别，一个手势只识别一次。这种设计通过多尺度协同检测与任务解耦，在保持实时性的同时显著提升检测精度，尤其针对重叠目标和小目标场景有显著优化效果[8]。

YOLOv5 提取图片中的车辆信息较多，采用中心检测的方式，哪个车辆最靠近像素点的中心位置，计算该车辆的车距。

2.2. 距离估算

摄像机的成像过程即透视投影，是从三维空间到二维空间的几何变换过程，而距离估计则是从二维空间转换到三维空间的过程，即逆透视投影。摄像机在成像时所有的平行直线会因为透视投影原理在图像上汇聚到一点，物体在图像中的形状和大小会发生扭曲，而逆透视变换(inverse perspective mapping, IPM)则是将原始图像转换[9]。采取如图 3 所示的标定版对相机进行标定。

创建 IPM 图需要三维坐标下的点 (X_w, Y_w, Z_w) 投影到像素坐标下 (u, v) 的关系，由几何关系可以将 IPM 的方程表示为 $(u, v, 1)^T = K \cdot T \cdot R(X_w, Y_w, Z_w, 1)^T$ ，其中 K 为旋转矩阵；为 T 变换矩阵；为 R 摄像头的参数矩阵。三个矩阵想乘得到，可以化为如下矩阵形式：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

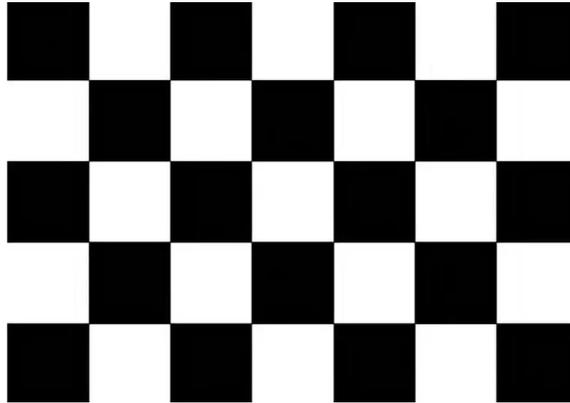


Figure 3. Camera calibration chart
图 3. 相机标定图

就可以寻找 IPM 图与原始图像的线性关系，通过标定得到比例系数以此求解距离。本文为了简化设计，假设车辆在较为平坦的道路，光线良好的条件下进行为了保证车距的计算较为精确，采用定位车牌的方式作为定位的参数，车牌使用像素提取的方式进行。根据标准标定图使用 Matlab 工具箱即可实现对摄像头的参数矩阵的获取[10]。

通过 YOLOv5 确定目标，通过像素提取的方法确定车牌位置信息，为逆透视变换法提供数据。逆透视变换法将斜视图转换为俯视图，能够消除透视畸变、还原路面本身信息，提高水平距离测量的准确性，并且不受场景的限制，可用于多种环境。

3. 实验设计与结果分析

3.1. 实验环境

数据集采用的是 UA-DETRAC 数据集，他是一个专门用于车辆检测和多目标跟踪的挑战性基准数据集，本文在数据集中选取适当的车辆信息，车辆信息的识别测距识别软件环境如表 1 所示。

Table 1. Experimental environmental parameters
表 1. 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. 实验结果

本文通过模型的训练得到模型，实验结果如图 4 所示，能够定位到当前车辆的位置，以及与当前车辆的距离，计算距离时已经将摄像头至车前头的距离从计算距离中移除。车辆定位的准确率与召回率如

表 2 所示，测距识别最大偏差小于 10%，数据详见表 3 测距结果统计。

Table 2. Algorithm result statistics

表 2. 算法结果统计

| 算法 | 准确率 | 召回率 |
|--------|-------|-------|
| YOLOv5 | 95.5% | 96.1% |

识别距离的偏差采用定点测距方式进行，将两辆车之间的具体，测试结果如表 3 所示。

Table 3. Algorithm result statistics

表 3. 测距结果统计

| 实际距离(m) | 测量距离(m) | 平均偏差 |
|---------|---------|--------|
| 1 | 0.92 | -7.87% |
| 2 | 1.97 | -1.64% |
| 3 | 3.19 | 6.33% |
| 4 | 3.81 | -4.63% |
| 5 | 5.25 | 5.10% |
| 6 | 6.14 | 2.35% |
| 7 | 6.80 | -2.79% |
| 8 | 7.30 | -8.75% |
| 9 | 8.52 | -5.33% |
| 10 | 10.60 | 5.97% |
| 11 | 11.30 | 2.69% |
| 12 | 11.50 | -4.14% |
| 13 | 14.23 | 9.49% |
| 14 | 14.54 | 3.88% |
| 15 | 15.06 | 0.39% |
| 16 | 14.51 | -9.29% |
| 17 | 17.28 | 1.65% |
| 18 | 17.74 | -1.45% |
| 19 | 18.59 | -2.17% |
| 20 | 18.98 | -5.09% |
| >20 | >20 | 3.32% |



Figure 4. Recognition results
图 4. 识别结果

4. 结语

在基于 YOLOv5 与车距检测算法的在一定程度上可以对车距进行识别，为车辆在行驶过程中的车辆判断起到辅助驾驶的作用，目前的算法只能识别到距离当前车辆最近的车距，而且车距信息的准确度可以通过深度学习等相关算法进一步提高，该方式也有一定的局限性，在路面相对平滑的区域有较好的识别效果，在道路比较崎岖的地方识别效果较差，该研究可以进一步得到提升。

参考文献

- [1] 骆绍焯, 彭展宏, 郭德清. 基于深度学习的车辆特征识别系统设计[J]. 鄂州大学学报, 2025, 32(1): 93-96.
- [2] 孙伟业, 刘云平. 基于改进目标匹配算法的车辆识别跟踪方法分析[J]. 集成电路应用, 2025, 42(3): 360-362.
- [3] 林海峰, 刘大鹏, 蔡慧. 改进 YOLOv5s 的交通摄像头视角下的车辆检测[J]. 计算机与数字工程, 2025, 53(5): 1327-1332+1398.
- [4] 梁秀满, 赵恒斌, 邵彭娟, 等. 基于改进 YOLOv5 的车辆检测方法[J]. 计算机应用与软件, 2025, 42(4): 217-222.
- [5] 刘丽丽, 王智文, 王亮, 等. 基于 BiFPN 和注意力机制改进 YOLOv5s 的车辆行人检测[J]. 现代电子技术, 2025, 48(3): 174-180.
- [6] 张照震. 基于单目视觉车距测量方法综述[J]. 汽车实用技术, 2022, 47(4): 153-157.
- [7] 刘大鹏, 蔡慧, 林海峰. 基于改进 YOLOv5s 的车辆检测研究[J]. 微型电脑应用, 2025, 41(1): 281-285+290.
- [8] 侯佩玉, 徐淼, 张明, 等. 基于 YOLOv5s 的自动驾驶车辆行人检测方法[J]. 北华大学学报(自然科学版), 2025, 26(1): 107-114.
- [9] 赵江洪, 刘蕊菱, 杨甲, 等. 几何单目视觉测距研究综述[J]. 测绘科学, 2023, 48(9): 49-65.
- [10] 杨勇超. 基于单目视觉的车距测量与碰撞预警方法研究[D]: [硕士学位论文]. 大连: 大连海事大学, 2023.