



计算机应用  
*Journal of Computer Applications*  
ISSN 1001-9081, CN 51-1307/TP

## 《计算机应用》网络首发论文

题目：基于轻量化 YOLOv4 的交通信息实时检测方法  
作者：郭克友, 李雪, 杨民  
收稿日期：2021-11-01  
网络首发日期：2022-04-20  
引用格式：郭克友, 李雪, 杨民. 基于轻量化 YOLOv4 的交通信息实时检测方法[J/OL]. 计算机应用. <https://kns.cnki.net/kcms/detail/51.1307.TP.20220418.1024.003.html>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于轻量化 YOLOv4 的交通信息实时检测方法

郭克友\*, 李雪, 杨民

(北京工商大学人工智能学院, 北京 100048)

(\*通信作者邮箱 guoky@th.btbu.edu.cn)

**摘 要:** 针对日常道路场景下的车辆目标检测问题, 提出了一种轻量化的 YOLOv4 交通信息实时检测方法。首先, 制作了一个多场景、多时段的车辆目标数据集, 并利用 K-means++ 对数据集进行预处理; 其次, 提出轻量化 YOLOv4 检测模型, 利用 MobileNet v3 替换 YOLOv4 的主干网络, 降低了模型的参数量, 引入深度可分离卷积以代替原网络中的普通卷积; 最后, 结合标签平滑和退火余弦算法, 新增 LeakyReLU 函数代替原有 MobileNet v3 浅层网络中的激活函数, 优化模型收敛效果。实验结果表明, 轻量化的 YOLOv4 权值文件为 56.4MB, 检测速率为 85.6FPS, 检测精度为 93.35%, 由此可见该方法可以为实际道路中的交通实时信息检测与应用提高参考。

**中图分类号:** 目标检测; 深度学习; 图像处理; 轻量化; YOLOv4

**中图分类号:** TP391.4

**文献标志码:** A

## Real-time detection method of traffic information based on lightweight YOLOv4

GUO Keyou\*, LI Xue, YANG Min

(School of Artificial Intelligence, Beijing Technology and Business University, Beijing 100048 China);

**Abstract:** Aiming at the problem of vehicle objection detection in daily road scenes, a real-time detection method of traffic information based on lightweight YOLOv4 was proposed. Firstly, a multi-scene and multi-period vehicle object dataset was made, which was preprocessed by K-means++. Secondly, a lightweight YOLOv4 detection model was proposed, and the backbone network of YOLOv4 was replaced by MobileNet V3, which reduced the number of parameters of the model. The depth separable convolution was introduced to replace the ordinary convolution in the original network. Finally, combined with label smoothing and annealing cosine algorithm, the LeakyReLU function was added to replace the activation function in the shallow network in the original MobileNet V3 to optimize the convergence effect of the model. The experimental results show that the weight file of lightweight YOLOv4 is 56.4MByte, the detection rate is 85.6 FPS (Frames Per Second), and the detection accuracy is 93.35%. It can be seen that this method can improve the reference for the detection and application of real-time traffic information in practical roads.

**Keywords:** object detection; deep learning; image processing; lightweight; YOLOv4

## 0 引言

随着技术的发展, 人们对交通监控系统的需求随之增加, 安全辅助驾驶系统的技术也逐渐走进大众的视野。车辆目标检测是安全辅助驾驶领域极具挑战的问题, 该领域主要分为传统目标检测算法和基于深度学习的检测算法。

传统目标检测方法主要分为三类, 以 MOG(Mixture Of Gaussians)<sup>[1]</sup>, MORG<sup>[2]</sup>和 GMG(Geometric Multi Grid)<sup>[3]</sup>为代表的背景差分法, 以两帧差分法<sup>[4]</sup>, 三帧差分法<sup>[5]</sup>为代表的帧差分法, 利用相邻两帧中对应的灰度保持不变原理评估二维图像

变化的光流场法<sup>[6]</sup>。然而, 传统目标检测算法存在鲁棒性差, 适应性弱等缺陷<sup>[7]</sup>。

近年来, 随着计算机技术的发展, 深度学习技术已在目标检测等任务取得了巨大的成就。基于深度学习的目标检测主要分为两大类<sup>[8]</sup>: (1) 基于区域建议的方法, 这类算法的典型代表是 R-CNN(Region Convolutional Neural Networks)<sup>[9]</sup>、SPP-net(Spatial Pyramid Pooling network)<sup>[10]</sup>、Fast R-CNN<sup>[11]</sup>、Faster R-CNN<sup>[12]</sup>等。(2) 基于无区域建议的方法, 这类方法主要采用回归的思想, 比较典型的算法如 YOLO(You Only Look Once)、SSD(Single Shot multi box

收稿日期: 2021-11-01; 修回日期: 2022-01-13; 录用日期: 2022-02-21。

**作者简介:** 郭克友(1975—), 男(汉), 黑龙江齐齐哈尔人, 副教授, 博士, 主要研究方向: 为图像处理、机电嵌入式; 李雪(1997—), 女(汉), 青海西宁人, 硕士研究生, 主要研究方向: SLAM、图像处理; 杨民(1996—), 男(汉), 四川达州人, 硕士研究生, 主要研究方向: 机器视觉、图像处理。

Detector)<sup>[13]</sup>。基于区域建议的方法精度高但速度慢,训练时间较长,基于无区域建议的方法准确度低,但速度快,对小物体的检测效果较差。

目前,应用最广泛的是 YOLO 系列检测算法。YOLOv1<sup>[14]</sup>的检测速度非常快,迁移能力强,但是对很小的群体检测效果不好。YOLOv2<sup>[15]</sup>经多种改进方法后,其检测速度都得到大幅提升,但检测精度提升效果并不明显。YOLOv3<sup>[16]</sup>中,作者提供了 Darknet-53 以及轻量级的 Tiny-Darknet 骨干网络,使用者可根据需求选择不同的骨干网络,其灵活性使得它在实际的工程中受到了大家的青睐。YOLOv4<sup>[17]</sup>在传统 YOLO 的基础上添加了许多实用技巧,实现了检测速度和检测精度的最佳权衡。YOLOv5 在性能上略弱于 YOLOv4,但在灵活性以及速度上都远强于 YOLOv4。

为应对复杂路况下,车辆多目标实时检测的挑战,本文选择检测精度和检测速度较为平衡的 YOLOv4 作为研究对象。为进一步提高模型的检测精度,并且使其能够部署在移动端,提出一种轻量化的 YOLOv4 交通信息实时检测方法。由于现有的大型公开数据集对道路复杂场景的适用性不强,本文通过针对真实交通场景构建了与之对应的数据集,并在测试前,对数据集进行 K-means++<sup>[18]</sup>聚类处理;除此之外,为提高检测性能,本文在 YOLOv4 的基础上对网络进行的修改,使得车辆目标的检测速率得到了大幅度的提高。

## 1 改进的 YOLOv4

YOLOv4 是 Alexey 继 YOLO、YOLOv2、YOLOv3 后的又一力作,其网络模型结构如下图 1 所示:它保留了 YOLOv3 的 head 部分,将主干网络修改为 CSPDarkent53(Cross Stage Partial Darknet53),同时采用 SSP (Spatial Pyramid Pooling) 空间金字塔池化的思想来扩大感受野, PANet (Path

Aggregation Network) 作为 neck 部分,并使用了多种训练技巧,对激活函数和损失函数进行了优化,使得 YOLOv4 不仅在检测精度方面有所提升,而且提高了检测的速度<sup>[19]</sup>。

CSPDarknet 作为 YOLOV4 算法的主干网络,经特征融合后,输出层的尺寸分别变为输入尺寸的  $\frac{1}{8}$ 、 $\frac{1}{16}$ 、 $\frac{1}{32}$ 。

CSPDarknet53 将原有的 Darknet53 中的残差块改为 CSP 网络结构。CSP-DarkNet 在每组 Residual block 加上一个 Cross Stage Partial 结构。并且, CSP-DarkNet 中也取消了 Bottleneck 的结构,减少了参数使其更容易训练。

SPP 以及 PANet 将主干网络提取的特征图进行融合, SPP 对全连接层前的卷积层进行不同池化大小的 pooling,然后拼接,由此增加网络感受野。PANet<sup>[20]</sup>提出了一种 bottom-up 的信息传播路径增强方式,通过对特征的反复提取实现了 featuremap1-featuremap3 的特征交互。PANet 可以准确地保存空间信息,有助于正确定位像素点,增强实例分割的过程。

最后, YOLOv4 Head 进行大小  $3 \times 3$  和  $1 \times 1$  两次卷积,对特征层的三个先验框进行目标及目标种类的判别,并进行非极大抑制处理和先验框调整,最后得到预测框。

为了进一步提高模型的检测速率,降低模型的参数量,本文对 YOLOv4 检测算法做如下改进:由 MobileNet-V3 成主干特征提取模块,经过特征提取后可以得到三个特征层:  $52 \times 52$ ,  $26 \times 26$ ,  $13 \times 13$ ,降低模型参数量;使用 LeakyReLU 激活函数代替 MobileNetv3 浅层网络中的 ReLU 激活函数,提高检测精度;利用深度卷积网络代替传统卷积网络。下面对网络结构的设计和参数设置进行阐述。

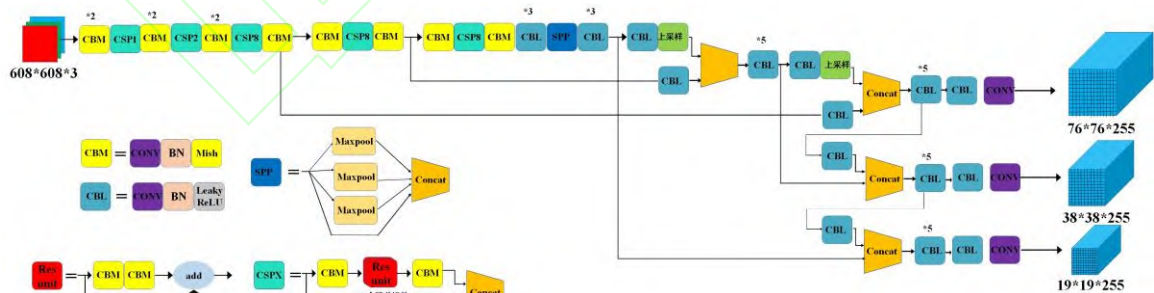


图1 YOLOv4 网络结构

Fig. 1 Structure of YOLOv4 network

### 1.1 主干特征提取网络

2019 年, Google 公司在 MobileNet 的基础上提出了 MobileNet v3 网络,在保持轻量化的同时, V3 模型进行了部分优化,经优化后的 MobileNet 模型衍生出 v3-Large 和 v3-Small 两个版本,在 image net 分类任务中, v3-Large 较 v2

模型分别提高了 3.2% 准确率以及 15% 检测速度, v3-Small 模型则分别提高了 4.6% 和 5%,在 coco 数据集的检测过程中,在满足精度的同时,检测速率提升了 25%<sup>[21]</sup>。MobileNet v3 的整体结构如图 2 所示:

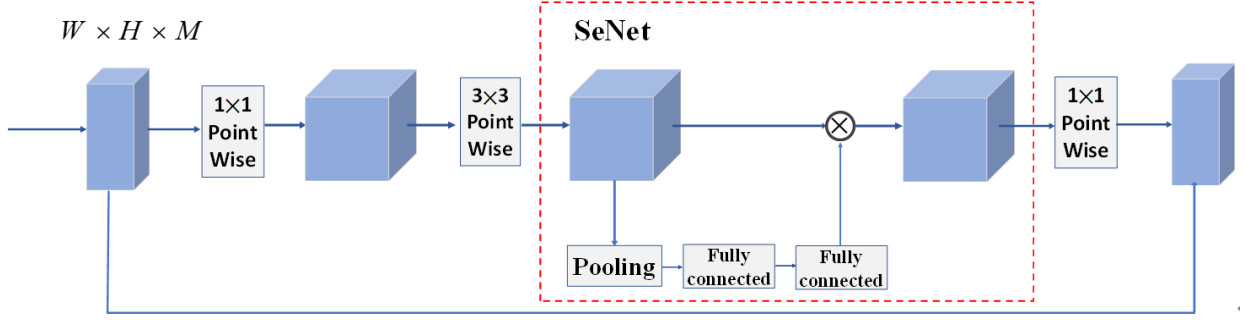


图2 MobileNet v3 网络结构

Fig. 2 MobileNet v3 network structure

MobileNet v3 在核心构架中引入了名为 Squeeze-and-Excitation 的神经网络, 该网络的核心思想是通过显示建模网络卷积特征通道之间的相互依赖关系来提高网络产生表示的质量, 即通过学习来自动获取到每个特征通道的重要程度, 并以此为结果提升有用的特征, 抑制用处较小的特征。本文将 MobileNet v3 网络代替传统 YOLOv4 的主干的网络, 减少参数量, 使模型更具轻量化。

## 1.2 深度可分离卷积

在深度可分离卷积出现之前, 神经网络基本由不同尺度的卷积核堆叠而成, 伴随技术的发展, 人们对网络模型的精度要求不断提高<sup>[22]</sup>。面对不断提高的深度核模型参数量, 聚焦于解决保证模型精度下如何降低参数量这一问题, 2017 年谷歌公司提出了基于深度可分离卷积的 MobileNet<sup>[23]</sup>, 随后提出了 MobileNet v2<sup>[24]</sup>, MobileNet v3<sup>[25]</sup> 系列网络。

MobileNet v3 使用深度可分离卷积代替一般卷积块, 深度可分离卷积 (DepthWise Conv) 由深度卷积和点卷积 (PointWise Conv) 组成, 由此极大减少模型的理论参数 (Params) 和每秒峰值速度 (FLOPs, Floating-point Operations Per Second)。

设输入数据为  $M \times M \times N$ , 卷积核大小为  $K \times K \times P$ , 设置步长为 1, 标准卷积操作利用卷积内核提取特征后将特征进行组合产生新的表示效果<sup>[26]</sup>。如图 3 所示, 此时, 标准卷积的参数量为:

$$W_{sc} = K \times K \times N \times P \quad (1)$$

计算量为:

$$O_{sc} = M \times M \times K \times K \times N \times P \quad (2)$$

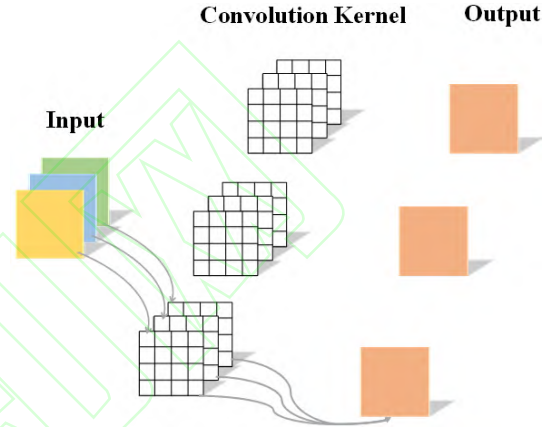


图3 传统卷积

FIG. 3 Traditional convolution

深度可分离卷积则是将提取特征和结合特征分为深度卷积和逐点卷积两步, 如图 4 所示, 在深度卷积过程中, 主要对每一个输入通道进行一个卷积核的操作, 随后利用  $1 \times 1$  的卷积将深度卷积的输出结果结合到特征中。随后通过  $1 \times 1$  卷积计算深度卷积的输出线性特征组合为新的特征, 深度卷积和  $1 \times 1$  卷积 (逐点卷积) 组成了深度可分离卷积<sup>[27]</sup>, 其参数量为:

$$W_{DSC} = K \times K \times N + N \times P \quad (3)$$

计算量为:

$$O_{DSC} = M \times M \times K \times K + M \times M \times N \times P \quad (4)$$

两种卷积对应参数比为:

$$F_w = \frac{W_{DSC}}{W_{sc}} = \frac{1}{P} + \frac{1}{K^2} \quad (5)$$

$$F_o = \frac{O_{DSC}}{O_{sc}} = \frac{1}{P} + \frac{1}{K^2} \quad (6)$$



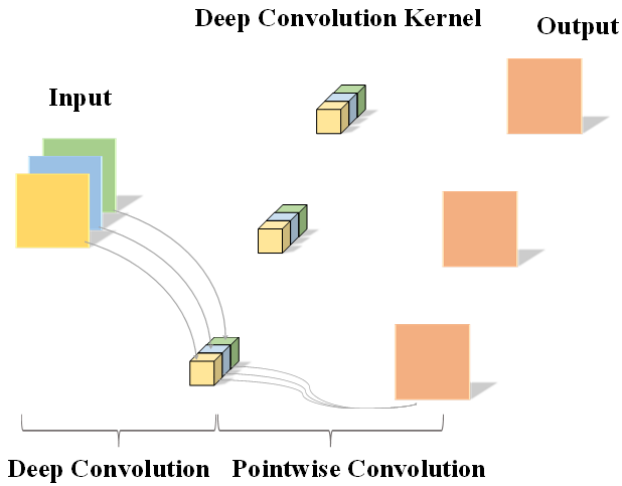


图4 深度可分离卷积

Fig. 4 Deep separable convolution

通过对比可知,深度可分离卷积与传统卷积相比更加高效,故本文用深度可分离卷积代替原始YOLOv4网络中的普通卷积,提高检测效率。

### 1.3 损失函数

MobileNet v3 浅层部分利用线性整流函数 ReLU (Rectified Linear Unit) 作为激活函数,该函数主要分为两部分,小于0的部分,激活函数输出为0;在大于0的部分,激活函数的输入为输入,其计算方法如下:

$$\text{ReLU} = \max\{0, z\} \quad (7)$$

该函数的收敛速度快,不存在饱和区间,在大于0的部分梯度固定为1,有效的解决 Sigmoid 中存在的梯度消失的问题;但是当个巨大的梯度经过 ReLU 神经元时,ReLU 函数将不具有激活功能,产生“Dead Neuron”现象。

为避免上述现象发生,本文中 MobileNet v3 特征提取网络采用 LeakyReLU 代替传统的 ReLU 函数,该函数可有效的解决神经元“死亡”问题。其计算方式如下:

$$\text{LeakyReLU}(x) = \begin{cases} x, & x \geq 0 \\ \alpha x, & x < 0 \end{cases} \quad (8)$$

实际中,LeakyReLU 的  $\alpha$  取值一般是 0.01。在使用过程中,对于 LeakyReLU 激活函数输入小于零的部分,也可以计算得到梯度,而不是像 ReLU 一样值为 0,由此可避免上述梯度方向锯齿问题。

## 2 数据集构建

### 2.1 数据集的制作

本实验采用的数据集由搭载 HP-F515 行车记录仪的实验平台,在北京市内真实道路场景下进行采集,包括城市主干道,桥梁公路,信号岔口等。采集流程为:为模拟驾驶员真

实驾驶情况,将行车记录仪搭载于后视镜处,行车路线全长 11 公里,行驶时间分别为 15:00-17:00, 19:00-21:00,该时间段内,测试路线车流量较大,存在部分拥堵路段。

提取录制的视频,对视频进行分帧,设置大小为  $1920 \times 1080$ ,并依据正常公路场景下的车辆目标类型,将车辆目标分为三种,如表 1 所示,分别是大型车(包括公交车、货车),中型车(包括越野车、面包车、平尾车等),小型车(包括轿车等)。实验共计 12614 张图片,并将标签分别设置为 2、3、4,如表 1 所示。

表1 三种目标及标签

Table 1 three objectives and labels

标签类别	标签类别描述	具体车辆类型
2	large vehicle	Buses, trucks, etc
3	medium	Off road vehicles, vans, etc
4	compact car	Car, etc

### 2.2 性能评价指标

目标检测中,常通过准确率(P)、召回率(R),计算平均精度(Average Precision, AP),平均AP值(mean Average Precision, mAP)作为目标检测的评价指标,如下式(9),式(10),式(11)所示。

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$R = \frac{TP}{TP + FN} \quad (10)$$

$$mAP = \frac{\sum AP}{N(class)} \quad (11)$$

其中:  $TP$  为检测正确的目标数量,  $FP$  为检测错误的目标数量,  $FN$  表示漏检数量,  $\sum AP$  表示所有缺陷类别的 AP 值总和,  $N(class)$  表示缺陷总类别数。

本文另选检测速度作为模型性能的评价指标之一,利用每秒可处理图像帧数量作为检测速度的评价指标。帧数越大,表示检测速度越快,可实现实时检测的要求,计算式(12)如下:

$$FPS = \frac{Num(\text{figure})}{TIME} \quad (12)$$

其中:  $Num(\text{figure})$  表述样本检测数量,  $TIME$  表示检测耗费的时间。

### 2.3 Anchor box 优化

为解决 K-means 算法由于随即产生的聚类中心,使得每次聚类效果不尽相同,进而影响模型的检测效果,为解决初始聚类中心不断变化等问题,本文采用 K-means++ 算法进行先验框的计算。设本文数据集内包含  $N$  个样本,数据集表示

为  $\alpha = \{x_i | x_i = (x_{i1}, x_{i2}, x_{i3}, L, x_{im}), i = 1, 2, LN\}$ , 取集合  $\alpha$  中  $k$  个数据作为初始的聚类中心点  $P = \{p_j | p_j = (p_{j1}, p_{j2}, p_{j3}, L, p_{jm}), j = 1, 2, Lk\}$ 。通过式 (9) 可得样本点  $x_i$  到质心  $p_j$  在  $m$  为空间的欧氏距离  $d$ , 随后依据式 (10) 遍历每个样本数据被选为下一个聚类中心的概率  $C^{[28]}$ 。

$$d(x_i, p_j) = \sqrt{\sum_{k=1}^m (x_{ik} - p_{jk})^2} \quad (13)$$

$$C(i) = \frac{d^2(x_i)}{\sum d^2(x_i)} \quad (14)$$

重复上述步骤, 直到选出  $k$  个聚类中心。分别计算样本点数据  $x_i$  到  $k$  个聚类中心的欧氏距离  $d$ , 按照邻近原则将样本点归类到距离其最近的聚类中心簇中, 取聚类中心簇内的样本点, 计算均值用来表示聚类中心, 经过迭代, 计算误差平方和, 直到  $I_{SSE}$  取最小值。其中  $I_{SSE}$  又称为畸变程度,  $I_{SSE}$  越小, 表明即便程度越低, 簇内样本点的关系越紧密。畸变程度会根据类别的增加而降低。

$$I_{SSE} = \sum_{j=1}^k \sum_{x_i \in \phi_j} d(x_i, p_j) \quad (15)$$

$$S_i = \frac{b(x_i) - a(x_i)}{\max\{b(x_i), a(x_i)\}} \quad (16)$$

$S_i$  为样本点  $x_i$  的轮廓系数,  $b(x_i)$  是  $x_i$  到聚类中心簇的平均距离,  $a(x_i)$  是  $b(x_i)$  中的最小值。

$$S_i = \frac{1}{n} \sum_{i=1}^N S_i \quad (17)$$

$S_i$  是样本点轮廓系数的平均值, 它反映了聚类的密集度和离散度, 由式 (17) 可知,  $S_i$  介于 0 和 1 之间,  $S_i$  越大, 表明类内样本的密集度越高, 各类样本的离散度越高, 聚类效果越好。

K-Means++ 算法通过轮廓系数和  $I_{SSE}$  的值得出最佳簇类数量, 使得分类效果更好。对自制数据集重新聚类后得到 9 组先验框为 (50.84, 43.92), (90.80, 61.82), (26, 87.25), (24.26, 49.68), (51.35, 189.15), (81.4, 57.78), (38.16, 38.13), (19.64, 21.95), (90.7, 124.8)。

### 3 实验与结果分析

#### 3.1 实验环境配置

实验采用的硬件配置为 Intel® Core i7-7700CPU @3.60GHz, 内存为 16GB, GPU 为 NVIDIA GeForce GTX1080Ti, 深度学习框架为 pytorch1.2.0 版本; Cuda 版本为 11.3; python 版本为 3.8.10。

本实验基于 Pytorch 深度学习框架, 数据集基于第三节的数据预处理, 训练过程为 350epoch, 前 50 个 epoch 利用冻结网络方式进行训练, 学习率为 0.001; 之后的 300 个 epoch 学习率为 0.0001, 训练过程采用了退火余弦算法以及标签平滑。

#### 3.2 改进模型有效性评价

为验证改进方法对 YOLOv4 模型的性能影响, 对上述四种改进方法进行消融实验设计, 以进行更为充分的对比, 由此证明所提方法的必要性和有效性。其中“√”表示在网络实验中使用了该改进策略, “—”表示未使用该策略。由表 2 可知, 将主干网络替换为 MobileNet v3 后, 权值文件由 161MB 变为了 53.7MB; 通过利用深度可分离卷积代替原始 YOLOv4 网络中的普通卷积, 网络结构中的参数得到了较大程度的减少; 最后, 轻量化后的模型与原模型比较可知, 虽然其减少了模型的参数量, mAP 降低了 0.76 个百分点, 由此可证明该方法具有一定的科学性和有效性。

表 2 消融实验结果对比表

Table 2 Comparison of ablation experiment

方法	MobileNet v3	深度可分离卷积	激活函数	K-means ++	mAP (%)	权值文件大小 (MB)	网络结构参数 (个)
YOLOv4	—	—	—	—	95	161	69,040,001
改进 1	√	—	—	—	95.87	53.7	39,989,933
改进 2	√	√	—	—	95.01	53.7	11,729,069
改进 3	√	√	√	—	94.03	53.7	11,729,069
改进 4	√	√	√	√	94.24	53.7	11,729,069

#### 3.3 定量性能评价

采用 AP、Recall、Precision 和 F1 作为算法性能评估指标, 并且与原始的 YOLOv4 网络进行对比, 各类别检测详细

结果如表 3 所示。YOLOv4 算法平均获得了 90.64% 的精确度, 91.6% 的召回率, 91.33% 的 F1, 轻量化的 YOLOv4 算法获得了 94.24% 的精确度, 90.71% 的召回率, 92% 的 F1 值。

表3 网络模型性能对比表

Table 3 Comparison of network model performance

标签类别	YOLO系列	AP (%)	Recall (%)	Precision (%)	F1 (%)
2	v4	99.24	90.91	90.91	91
	Tiny v4	90.06	87.16	89.58	88
	x	97.52	94.33	95.00	95
	轻量化	<b>94.98</b>	<b>90.87</b>	<b>90.26</b>	<b>91</b>
3	v4	92.27	92.54	92.54	93
	Tiny v4	90.08	86.75	90.05	88
	x	95.62	91.62	94.47	93
	轻量化	<b>93.82</b>	<b>90.97</b>	<b>94.71</b>	<b>93</b>
4	v4	93.06	91.35	88.48	90
	Tiny v4	88.87	84.33	89.51	87
	x	94.28	90.43	92.91	92
	轻量化	<b>93.91</b>	<b>90.31</b>	<b>93.73</b>	<b>92</b>

网络的评价指标除了 mAP 以外, 还有检测速率的判别依据 FPS (Frames Per Second)。为了全面的和原始 YOLOv4 目标检测算法进行差异对比, 结果如表 4 所示, 可看出, 当输入分辨率一致时, 轻量化的 YOLOv4 检测算法在实时性上有所上升, 虽然 mAP 下降了 0.76%, 但是检测速率提升了 79%, 不仅可以减少误检、漏检还能保证检测精度, 提高了检测速度。

表4 网络模型检测速度对比

Table 4 Comparison of network model detection speed

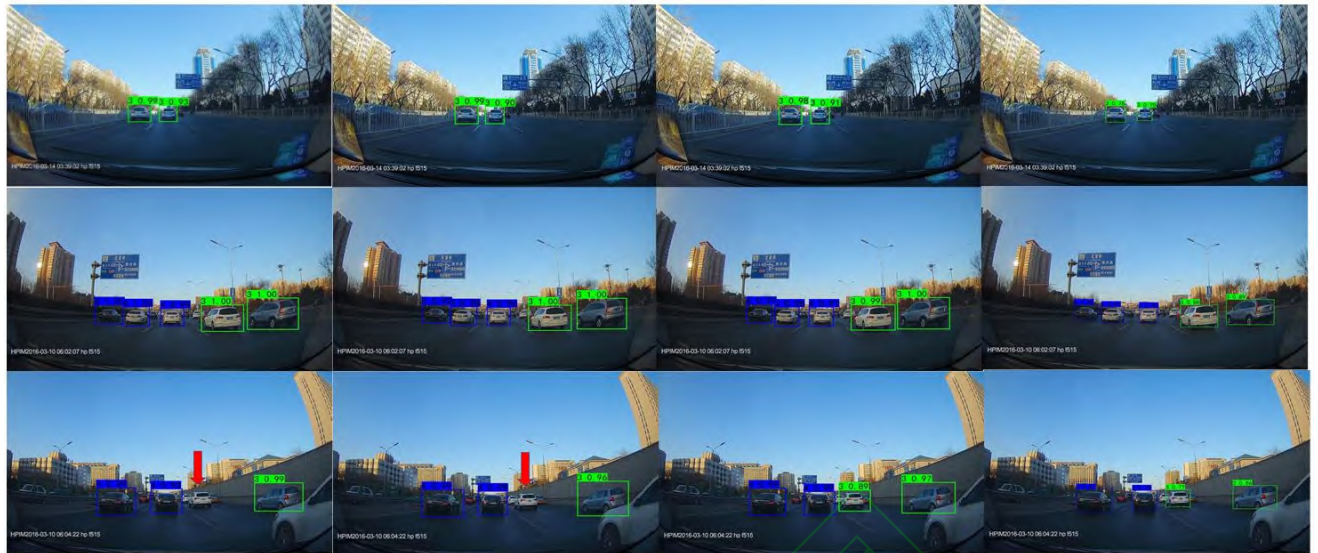
网络名称	主干网络	检测速度	输入尺寸	mAP
YOLOv4	CSPDarknet 53	47.74	416×416	95
Tiny YOLOv4	CSPDarknet 53-tiny	197	416×416	89.3
YOLOx	CSPDarknet	100.2	640×640	95.8
轻量化 YOLOv4	MobileNet v3	85.6	416×416	94.24

本文使用 MobileNet-v3 和深度可分离卷积代替 YOLOv4 的主干网络 and 标准卷积使得计算量下降到  $1.1 \times 10^7$ , 使得检

测速度由原始 47.74FPS 升高至 85.6FPS, 从以上分析可知, 轻量化的 YOLOv4 检测算法比 YOLOv4 检测算法在保证检测精度的同时, 提高了检测效率; 但与同样为轻量化系列的 Tiny YOLOv4 以及 YOLOx 相比在检测速度方面有较大的差距。

### 3.4 检测结果对比图

为了进一步直观的展示本文算法的有效性, 图 5 至图 7 展示了原始 YOLO 系列检测算法和轻量化的 YOLOv4 检测算法在车辆目标数据集的检测结果。通过图 5 的对比, 在道路情况较为简单的情况下, 三类检测器对车辆目标的检测效果良好; 当道路中出现较多车辆, 且车辆间不存在大量重叠和遮挡的情况下, 检测器的检测效果较为统一, 均可检测出道路中的中型车以及小型车; 当道路中出现不完整车辆目标、遮挡目标和小目标时, 原始 YOLOv4 检测模型、Tiny YOLOv4 以及 YOLOx 模型的检测显示, 它们对远处目标的检测效果略差于轻量化 YOLOv4 检测效果。



(a) 原始 YOLOv4 检测器 (b) Tiny YOLOv4 检测器 (c) 轻量化 YOLOv4 检测器 (d) YOLOx 检测器

图 5 检测效果对比

Fig. 5 Comparison of detection results

为了进一步对比三种模型算法针对不完整目标和小目标的检测能力,选择 100 张含有大量重叠车辆目标以及小目标的图片作为测试数据集,其检测性能的对比如图 6 所示,

随后进行更直观的对比实验,选择存在重叠目标的数据集进行车辆目标检测,检测结果如图 6 所示。

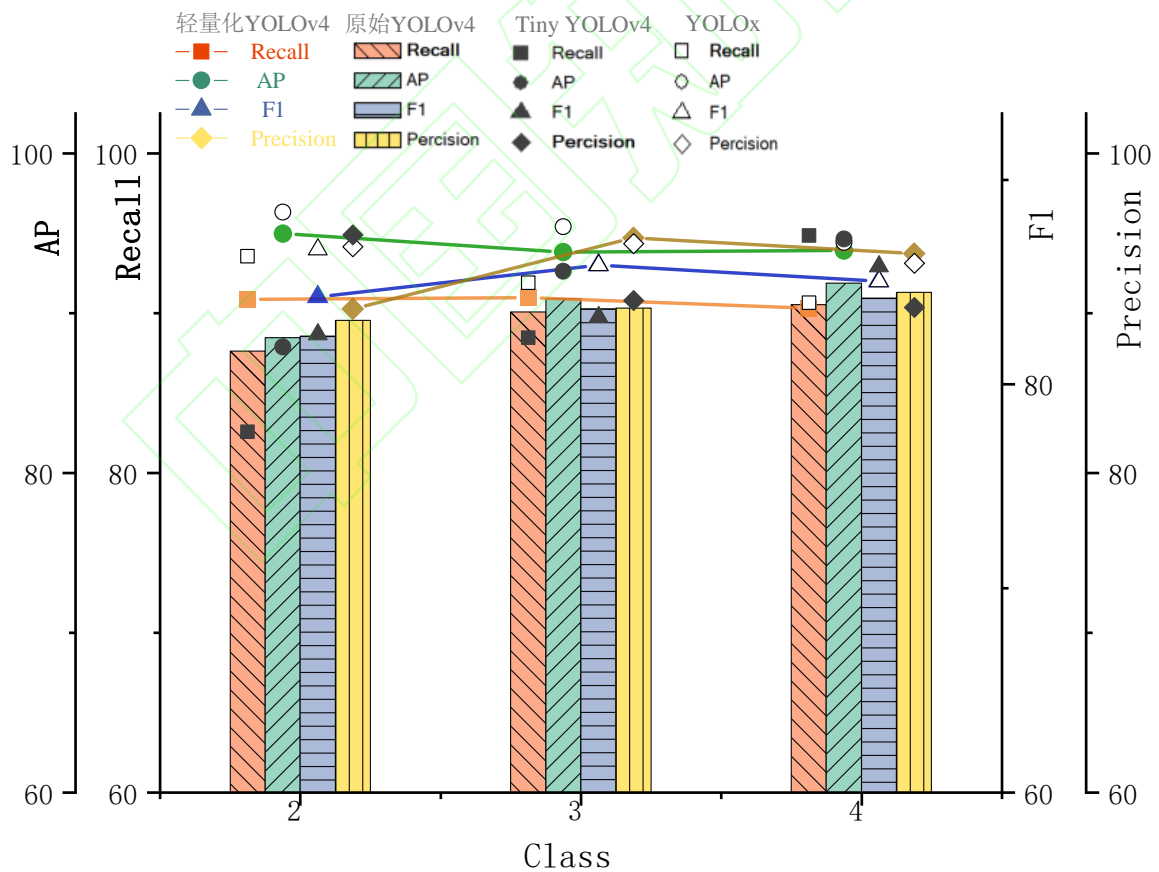


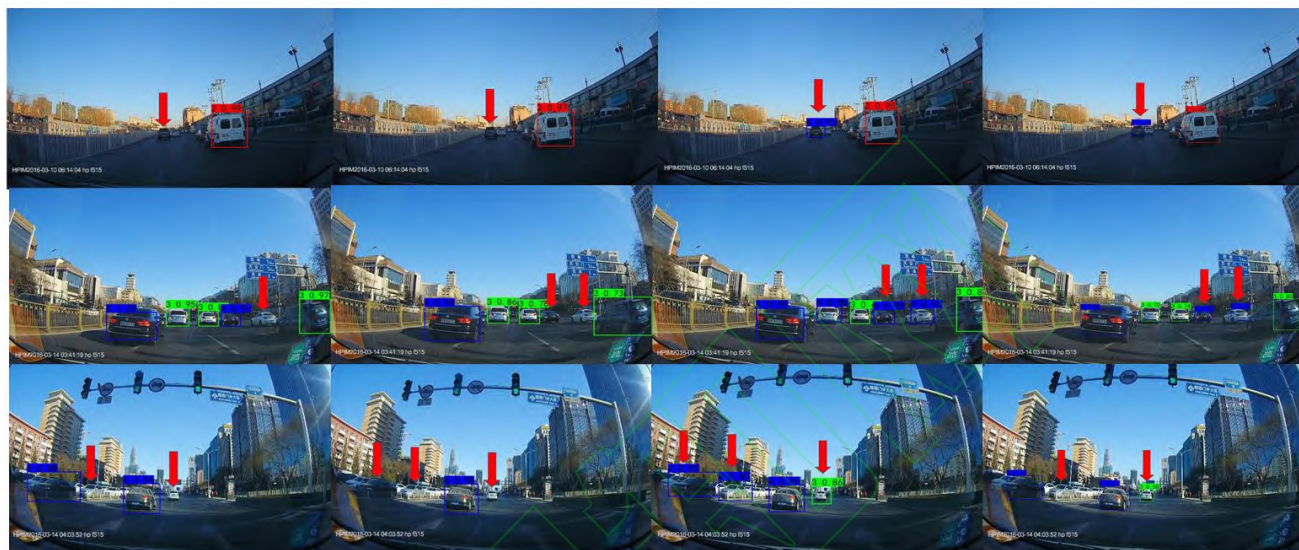
图 6 四种模型的评价指标对比

Fig. 6 Comparison of evaluation indexes of the four models



通过图 7 的对比, 面对道路较为简单的情况, 即道路中车辆较少, 且无明显重叠现象时, 原始 YOLOv4 检测算法可以检测出 1 个大型车, 无法识别远处的车辆小目标, Tiny YOLOv4 无法检测出远处车辆小目标, 轻量化的 YOLOv4 检测算法共测出 2 个车辆目标, 包含远处的小目标车辆, YOLOx 检测器也可检测出 2 个车辆目标, 但其检测精确度不如轻量化的 YOLOv4; 当车辆目标数量增多, 但车辆之间的重叠情况较少时, 远处的小目标可以在轻量化的 YOLOv4

检测模型中被测得, 但是原始 YOLOv4、Tiny YOLOv4 以及 YOLOx 都存在远处的重叠目标的漏检现象, 其中 Tiny YOLOv4 检测器漏检情况较原始 YOLOv4 检测器较为严重; 当道路中出现栅栏等遮挡物且车辆目标较多的情况下, 原始 YOLOv4 检测器检测出了 2 个车辆小目标, Tiny YOLOv4 检测器仅检测出 1 个车辆小目标, YOLOx 检测出了 3 个车辆目标, 面对遮挡目标存在一定的漏检情况, 轻量化的 YOLOv4 检测器检测出了 3 个近处目标以及 1 个远处小目标。



(a) 原始 YOLOv4 检测器

(b) Tiny YOLOv4 检测器

(c) 轻量化 YOLOv4 检测器

(d) YOLOx 检测器

图 7 检测效果对比

Fig. 7 Comparison of detection results

综合上述分析可知, 轻量化的 YOLOv4 模型和其他模型都可以准确的识别出近处的未被遮挡的完整目标, 但是在部分复杂路况下, 原始的 YOLOv4 无法识别出近处的被遮挡目标, Tiny YOLOv4 对近处的不完整目标存在一定的漏检, 对于不完整的目标、重叠目标以及远处的小目标的漏检情况较为严重, YOLOx 面对远处遮挡小目标存在一定的遮挡现象。

## 4 结 论

本文提出了一种轻量化的 YOLOv4 交通信息实时检测算法, 为模拟真实道路路况信息, 利用行车记录仪录制不同路段的北京道路场景信息并制作相应数据集, 随后利用 K-Means++ 算法对锚框进行聚类。使用 MobileNet-v3 网络代替 YOLOv4 主干网络, 并在 MobileNet-v3 网络中的浅层激活函数替换为 LeakyReLU 激活函数, 最后将 YOLOv4 中的标准卷积替换为深度可分离卷积降低运算量。实验表明: 轻量化的 YOLOv4 检测算法与传统 YOLOv4 检测算法相比, 在提高检测速率为 85.6FPS, mAP 值为 91.94%, 本模型可以为 Tiny YOLOv4 检测算法进行针对复杂场景下的重叠目标、不完整目标和小目标提供辅助检测, 但本模型仍有改进空间,

如何提高检测速率, 并使用与更丰富的检测场景是接下来待解决的问题。

## 参考文献

- [1] KAEWTRAKULPONG P, BOWDEN R. An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection[C]// Proceedings of 2nd European Workshop on Advanced Video-Based Surveillance Systems, Cham: Springer 2002:135-144
- [2] ZIVKOVIC Z, HEIJDEN F V D. Efficient adaptive density estimation per image pixel for the task of background subtraction[J]. Pattern Recognition Letters. 2006(27): 773-780.
- [3] GODBEHERE A, MATSUKAWA A, GOLDBERG K. Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation[C]// Proceedings of the 2012 American Control Conference. Piscataway: IEEE, 2012: 4305-4312.
- [4] SCHARSTEIN D, SZELISKI R. A Taxonomy And Evaluation Of Dense Two-Frame Stereo Correspondence Algorithms[J]. International Journal of Computer Vision - IJCV. 2000,47(1): 7-42.
- [5] BERGEN J, BURT P, HINGORANI R, et al. A three-frame algorithm for estimating two-component image motion[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 1992,14(9): 886-896.
- [6] BROX T, BRUHN A, PAPENBERG N, et al. High Accuracy Optical Flow Estimation Based on a Theory for Warping[C] // Proceedings of the 2012 European Conference on Computer Vision, Cham: Springer, 2004: 25-36.

- [7] 王滢喧, 宋焕生, 梁浩翔, 等. 基于改进的 YOLOv4 高速公路车辆目标检测研究[J]. 计算机工程与应用. 2021, 57(13): 218-226. (WANG Y X, SONG H S, LIANG H X, et al. Highway Vehicle Object Detection Based on Improved YOLOv4 Method [J] Computer engineering and applications 2021, 57(13): 218-226.)
- [8] 张慧, 王坤峰, 王飞跃. 深度学习在目标视觉检测中的应用进展与展望[J]. 自动化学报. 2017, 43(8): 1289-1305. (ZHANG H, WANG K F, WANG F Y. Advances and Perspectives on Applications of Deep Learning in Visual Object Detection[J]. Journal of automation 2017, 43(8): 1289-1305.)
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. // Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2014:580-587.
- [10] HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2014,37(9): 1904-1915.
- [11] GIRSHICK R. Fast r-cnn[C]. // Proceedings of the 2015 IEEE International Conference on Computer Vision, Piscataway: IEEE, 2015:1440-1448.
- [12] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2016,39(6): 1-10.
- [13] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]. // Proceedings of the 14th European Conference, Amsterdam, Cham: Springer, 2016: 21-37.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]. // Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2016: 779-788.
- [15] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2017:6517-6525.
- [16] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement [EB/OL]. [2018-04-18] <https://arxiv.org/pdf/1804.02767.pdf>.
- [17] BOCHKOVSKIY A, WANG C, LIAO H. YOLOv4: Optimal Speed and Accuracy of Object Detection [EB/OL] [2020-4-23] <https://arxiv.org/pdf/2004.10934.pdf>.
- [18] MCQUEEN J. Some methods for classification and analysis of multivariate observations[J]. Computer and Chemistry. 1967,5(1): 257-272.
- [19] 钟志峰, 夏一帆, 周冬平, 等. 基于改进 YOLOv4 的轻量化目标检测算法 [J/OL]. 计算机应用 : 1-8[2021-9-29]. <https://kns.cnki.net/kcms/detail/51.1307.TP.20210929.1334.012.html> (ZHONG Z F, XIA Y F, ZHOU D P, et al. Lightweight object detection algorithm based on improved YOLOv4[J/OL]. Journal of Computer Applications:1-8.[2021-9-29].<https://kns.cnki.net/kcms/detail/51.1307.TP.20210929.1334.012.html>)
- [20] LIU S, QI L, QIN H, et al. Path Aggregation Network for Instance Segmentation[C]. // Proceeding of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2018: 8759-8768.
- [21] 叶子勋, 张红英. YOLOv4 口罩检测算法的轻量化改进[J]. 计算机工程与应用. 2021,57(17):1-19. (YE Z X, ZHANG H Y. Lightweight Improvement of YOLOv4 Mask Detection Algorithm[J]. Computer Engineering and Applications. 2021,57(17): 1-19.)
- [22] 邓若辰, 彭程, 边赟. 基于深度可分离卷积的指静脉识别算法[J]. 计算机应用. 2020, 40(S2): 199-203. (DENG R C, PENG C, BIAN Y. Finger Vein recognition algorithm based on depthwise convolution[J]. Journal of Computer Applications. 2020(40(S2)): 199-203.)
- [23] HOWARD A, ZHU M, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. [EB/OL]. [2017-04-17] <https://arxiv.org/pdf/1704.04861.pdf>
- [24] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks [C]. //Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2018: 4510-4520.
- [25] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[C]. // Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision, Piscataway: IEEE,2019:1314-1324.
- [26] KAMIBABA R, KONISHI K, FUKADA Y, et al, et al. Next generation 650V CSTBTM with improved SOA fabricated by an advanced thin wafer technology[C]. // Proceedings of the 2015 IEEE 27th International Symposium on Power Semiconductor Devices & IC's (ISPSD), Piscataway: IEEE, 2015:29-32
- [27] 王帅, 彭意兵, 何顶新. 基于深度可分离卷积神经网络的关键词识别系统[J]. 微电子学与计算机. 2019, 36(9): 103-108. (WANG S, PENG Y B, HE X X. Keywords spotting system based on deepwise separable convolutional neural network [J]. Microelectronics & Computer. 2019, 36(9): 103-108.)
- [28] 李维刚, 杨潮, 蒋林, 等. 基于改进 YOLOv4 算法的室内场景目标检测[J/OL]. 激光与光电子学进展: 1-19[2021-8-3]. <https://kns.cnki.net/kcms/detail/31.1690.TN.20210802.1726.044.html> (LI W G, YANG C, JIANG L, et al. Indoor Scene Object Detection Based on Improved YOLOv4 Algorithm [J/OL]. Laser & OptoelectronicsProgress:1-19[2021-8-3]<https://kns.cnki.net/kcms/detail/31.1690.TN.20210802.1726.044.html>)

GUO Keyou, born in 1975, ph. D, associate professor, his main research interests: image processing, electromechanical embedded.

LI Xue, born in 1997, M. S. candidate, her main research interests are SLAM and image processing.

YANG Min, born in 1996, M. S. candidate, his main research interests are machine vision and image processing.