

## Tiny YOLOV3目标检测改进

马立, 巩笑天, 欧阳航空

引用本文:

马立, 巩笑天, 欧阳航空. Tiny YOLOV3目标检测改进[J]. 光学精密工程, 2020, 28(4): 988–995.

MA Li, GONG Xiao-tian, OUYANG Hang-kong. Improvement of Tiny YOLOV3 target detection[J]. *Optics and Precision Engineering*, 2020, 28(4): 988–995.

在线阅读 View online: <https://doi.org/10.3788/OPE.20202804.0988>

## 您可能感兴趣的其他文章

Articles you may be interested in

### 基于改进YOLOv2的快速安全帽佩戴情况检测

Fast helmet-wearing-condition detection based on improved YOLOv2

光学精密工程. 2019, 27(5): 1196–1205 <https://doi.org/10.3788/OPE.20192705.1196>

### 采用深度级联卷积神经网络的三维点云识别与分割

Recognition and segmentation of three-dimensional point cloud based on deep cascade convolutional neural network

光学精密工程. 2020, 28(5): 1187–1199 <https://doi.org/10.3788/OPE.20202805.1187>

### 窗口融合特征对比度的光学遥感目标检测

Optical remote sensing object detection based on fused feature contrast of subwindows

光学精密工程. 2016, 24(8): 2067–2077 <https://doi.org/10.3788/OPE.20162408.2067>

### 面向激光跟踪仪跟踪恢复的合作目标视觉检测

Visual detection of targetball for laser tracker target tracking recovery

光学精密工程. 2020, 28(2): 271–282 <https://doi.org/10.3788/OPE.20202802.0271>

### 构建多尺度深度卷积神经网络行为识别模型

Action recognition model construction based on multi-scale deep convolution neural network

光学精密工程. 2017, 25(3): 799–805 <https://doi.org/10.3788/OPE.20172503.0799>

文章编号 1004-924X(2020)04-0988-08

# Tiny YOLOV3 目标检测改进

马 立, 巩笑天, 欧阳航空\*

(上海大学 机电工程与自动化学院, 上海 200444)

**摘要:**针对 Tiny YOLOV3 目标检测算法在实时检测中对行人等小目标漏检率高的问题, 对该算法的特征提取网络、预测网络、损失函数等进行研究改进。首先, 在特征提取网络中增加 2 步长的卷积层, 代替原网络中的最大池化层进行下采样; 接着, 使用深度可分离卷积构造反残差块替换传统卷积, 降低模型尺寸和参数量, 增加高维特征提取; 然后, 在原网络两尺度预测的基础上增加一尺度, 形成三尺度预测; 最后, 对损失函数中的边界框位置误差项进行优化。实验结果表明, 改进后的 Tiny YOLOV3 算法的目标检测准确率比原算法提高了 9.8%, 满足实时性要求, 具有一定鲁棒性。本文方法能够更好地提取目标特征, 多尺度预测和边界框位置误差的改进能更准确地对目标进行检测。

**关键词:**目标检测; Tiny YOLOV3; 深度可分离卷积; 反残差块; 多尺度预测

**中图分类号:** TP391 **文献标识码:** A **doi:** 10.3788/OPE.20202804.0988

## Improvement of Tiny YOLOV3 target detection

MA Li, GONG Xiao-tian, OUYANG Hang-kong\*

(School of Mechatronic Engineering and Automation,  
Shanghai University, Shanghai 200444, China)

\* Corresponding author, E-mail: hankow@shu.edu.cn

**Abstract:** The Tiny YOLOV3 target detection algorithm has a high error rate for small targets, such as pedestrians, in real-time detection. Therefore, this study aimed to improve the feature extraction network, prediction network, and loss function of the algorithm. First, a two-step convolution layer was added to the feature extraction network to replace the maximum pooling layer in the original network for downsampling. Second, the traditional convolution was replaced with an anti-residual block constructed by a deep convolutional convolution to reduce the model size as well as number of parameters and increase the high-dimensional feature extraction. Third, based on the original two-scale prediction of the network, a scale was added to form a three-scale prediction. Finally, the boundary box position error in the loss function was optimized. The experimental results demonstrate that the improved Tiny YOLOV3 algorithm achieves a target detection accuracy that is 9.8% higher than the original algorithm, satisfied the real-time requirement, and demonstrated robustness. The proposed method can better extract target features, and the multi-scale prediction and improvement of the

收稿日期: 2019-11-98; 修订日期: 2020-01-01.

基金项目: 国家自然科学基金资助项目 (No. 61573238); 国家重点研发计划资助项目 (No. 2018YFB1309200)

boundary box position error can detect targets more accurately.

**Key words:** target detection; Tiny You Only Look Once V3(YOLOV3); depth separable convolution; anti-residual block; multi-scale prediction

## 1 引言

随着社会发展和科技进步,无人驾驶车辆成为未来交通的一大趋势。目标检测方法是无人驾驶系统必不可少的部分,可分为浅层学习和深度学习两种方法。浅层学习一般根据模板和图像稳定的特征点获得模板与场景中对象的对应关系来检测目标,使用 AdaBoost<sup>[1]</sup> 和 Canny<sup>[2]</sup> 等特征提取方法和粒子群 SVM(Support Vector Machine)<sup>[3]</sup> 等分类方法进行检测。近年来,随着硬件运算能力提高,基于卷积神经网络的深度学习<sup>[4]</sup> 得到迅速发展。相比浅层学习,深度学习能够更好地提取特征,目前分为两种:一种是结合预测框和卷积神经网络分类,为二阶段算法,以 R-CNN<sup>[5-7]</sup> 系列为主要发展方向,虽然检测精度大幅提升,但是速度非常慢;另一种是将分类和预测合并为一步,为单阶段算法,以 SSD(Single Shot MultiBox Detector)<sup>[8]</sup> 系列和 YOLO(You Only Look Once)<sup>[9-11]</sup> 系列为主,通过卷积神经网络提取特征,直接预测边界框坐标和类别。YOLOV3 是 YOLO 系列最新的目标检测算法,融合了特征金字塔<sup>[12]</sup>,借鉴了残差网络<sup>[13]</sup> 和多尺度预测网络<sup>[14-15]</sup>,大幅提升了目标检测速度和准确率。但这些算法模型较大,对嵌入式系统来说,无法满足实时性要求,不适用于无人驾驶实际场景。

Tiny YOLOV3 是 YOLOV3 的简化版本,卷积层数大幅减少,模型结构简单,不需占用大量内存,是目前最快的目标实时检测算法,但检测精度较低,尤其是在行人等小目标检测<sup>[16-17]</sup> 中,漏检率较高。本文针对这一问题展开研究,在特征提取网络阶段添加 2 步长的卷积层,深化的网络可以更好地提取特征,提高检测精度,但增加卷积层数,会导致参数量剧增,这将极大地增加计算量并占用内存资源,因此本文采用深度可分离卷积构造反残差块,在提高检测精度

的同时实现快速计算以满足实时检测要求。改进预测网络和损失函数,实现三尺度预测,进一步提高检测精度。

## 2 Tiny YOLOV3 算法改进

### 2.1 Tiny YOLOV3 算法

Tiny YOLOV3 算法是在 YOLOv3 算法基础上应用于嵌入式平台的轻量级实时检测算法,在检测精度上有所降低,但是实现了模型压缩,Tiny YOLOV3 将 YOLOv3 特征检测网络 darknet-53 缩减为 7 层传统卷积和 6 层 Max Pooling(最大池化)层,采用  $13 \times 13$ 、 $26 \times 26$  两尺度预测网络对目标进行预测,网络结构如图 1 所示。

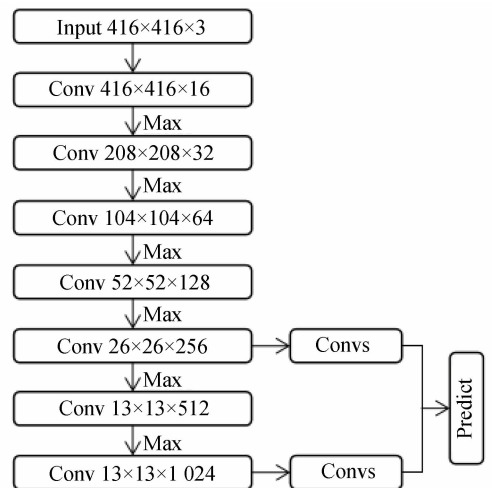


图 1 Tiny YOLOV3 网络结构

Fig. 1 Tiny YOLOV3 network structure

Tiny YOLOv3 将输入图像划分为  $S \times S$  的网格,在每个网格内预测  $B$  个边界框,对  $C$  类目标进行检测,输出每类目标的边界框和置信度。置信度由每个网格中包含检测目标的概率和输出边界框的准确度共同确定,其公式为:

$$Confidence = P_r(object) * IOU_{pre}^{truth}, \quad (1)$$

其中:  $P_r(object)$  为目标存在概率,  $IOU_{pre}^{truth}$  为边界框与真实框的交并比。

Tiny YOLOv3 的损失函数 Loss 主要从边界框位置误差项、边界框的置信度误差项以及分类预测误差项 3 个方面定义：

$$\begin{aligned}
 Loss = & \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \\
 & \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] + \\
 & \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 + \\
 & \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{obj} (C_i - \hat{C}_i)^2 + \\
 & \sum_{i=0}^{s^2} l_i^{obj} \sum_{(c \in \text{classes})} (P_i(c) - \hat{P}_i(c))^2, \quad (2)
 \end{aligned}$$

其中：第 1、2 项为位置坐标误差项，第 3、4 项为置信度误差项，第 5 项为分类预测误差项。

### 2.2 构造反残差块

传统卷积的参数计算量随着卷积层数增加而大幅增长，本文使用深度可分离卷积代替传统卷积，将传统卷积转变为深度卷积和逐点卷积两部分，可有效减少模型大小的参数计算量。但这种方法会随着卷积层数增加而出现梯度消失问题，残差结构能够解决这个问题，但残差结构会压缩特征图损害特征表达。本文在特征提取过程中采用反残差块，先通过  $1 \times 1$  卷积扩张通道，然后用  $3 \times 3$  的深度卷积提取高维空间特征，最后再通过  $1 \times 1$  的点卷积层将深度卷积结果映射到新的通道空间。残差块与反残差块结构如图 2 所示。图中： $n$  为输入通道数， $t$  为扩充或压缩通道的倍数， $C$  为通道数。本文反残差块参数见表 1。表 1 中： $h, w$  分别为特征图的高和宽； $n$  为特征图的通道数； $s$  为步长。

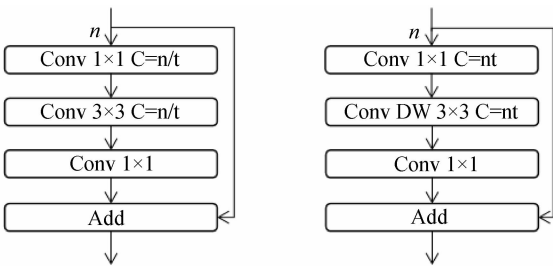


图 2 残差块与反残差块结构

Fig. 2 Figure 2 residual block and anti-residual block structure

表 1 反残差块参数

Tab. 1 Anti-residual block parameter

输入	卷积	输出
$h \times w \times n$	Conv $1 \times 1$ , Relu	$h \times w \times 2n$
$h \times w \times 2n$	Conv DW $3 \times 3$ , Relu	$h/s \times w/s \times 2n$
$h/s \times w/s \times 2n$	Conv $1 \times 1$	$h/s \times w/s \times 2n$

### 2.3 网络模型改进

为了解决 Tiny YOLOv3 对行人等小目标检测精度低，漏检率高的问题，本文在原有网络基础上进行改进，网络结构如图 3 所示，其中，ARB 为反残差块 (Anti-Residual Block)，虚线框中为网络特征提取部分。

在特征提取网络中，通过增加卷积层提高特征提取量，用步长为 2 的卷积代替原网络中的最大池化层进行下采样，并用深度可分离卷积构造的反残差块代替传统卷积。改进的网络由 12 个反残差块构成，通过反残差块，扩张特征图通道提取高维特征，再进行通道降维，得到特征图。在增加特征提取的同时，有效降低模型尺寸和参数计算量。同时，在原网络  $26 \times 26, 13 \times 13$  两尺度预测目标的基础上增加一上采样层 upsample，形成  $52 \times 52, 26 \times 26, 13 \times 13$  三尺度预测，进一步提高目标检测准确率。

表 2 列出了 YOLOV3, Tiny YOLOV3 及本文改进的网络模型大小及处理一张图片所需要的计算量。可以看出，本文改进后的网络模型与 Tiny YOLOV3 相比，模型尺寸仅大 0.4 M，处理一张图片所需计算量增加 0.17GFLOPs，远小于 YOLOV3 模型计算量，在模型大小和运算量上极具优势，满足嵌入式系统实时检测要求。

表 2 不同网络模型大小及计算量

Tab. 2 Different network model sizes and calculations

算法名称	模型尺寸/M	浮点计算量/GFLOPs
YOLOV3	246.5	65.7
Tiny YOLOV3	34.7	5.56
改进算法	35.1	5.73

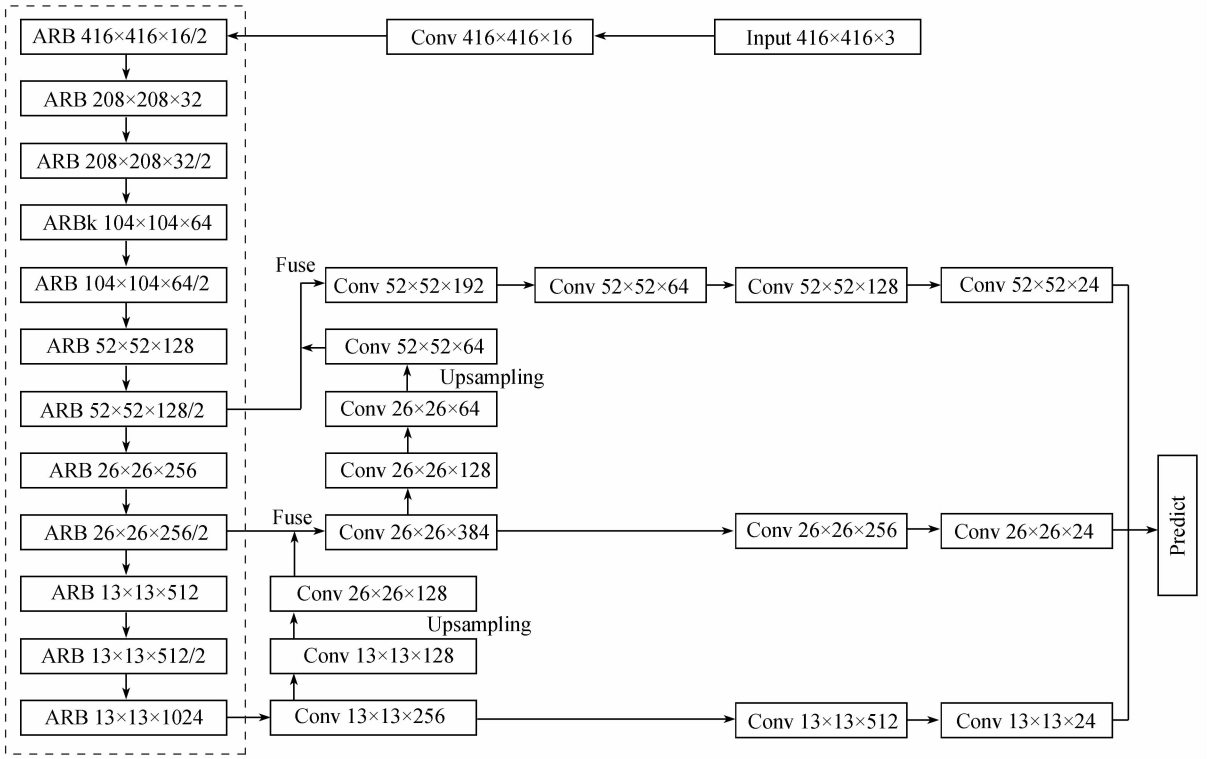


图3 改进的 Tiny YOLOV3 网络模型

Fig. 3 Improved Tiny YOLOV3 network model

## 2.4 损失函数改进

交并比 (Intersection Over Union, IOU) 度量预测框和真实框之间的交集, 是在目标检测基准中使用的最流行的评估方法, 但是在预测框和真实框无交集的情况下不能进行度量评估。为此, 本文对损失函数中边界框位置误差项进行优化, 采用广义交并比 (Generalized Intersection Over Union, GIOU) 解决这个问题。GIOU 为 IOU 与不含最小封闭面的预测框真实框交集与最小封闭面的比值的差。交并比和广义交并比的定义如式(3)和式(4):

$$IOU = \frac{|A \cap B|}{|A \cup B|}, \quad (3)$$

$$GIOU = IOU - \frac{|C \setminus (A \cap B)|}{|C|}, \quad (4)$$

其中:  $A$  表示真实框,  $B$  表示预测框,  $C$  表示预测框和真实框之间的最小封闭面。

使用 GIOU 检测目标时, 比较两个轴对齐的边界框, 真实框和预测框组成的交叉面和最小封闭面均为矩形, 可用解析解来计算。

假设:  $M^p = (x_1^p, y_1^p, x_2^p, y_2^p)$  为边界框的预测值,  $M^g = (x_1^g, y_1^g, x_2^g, y_2^g)$  为真实值, 且  $(x_2^p > x_1^p,$

$y_2^p > y_1^p)$ , 则对应的面积  $N^p$  和  $N^g$  为式(5)和式(6):

$$N^p = (x_2^p - x_1^p) \times (y_2^p - y_1^p), \quad (5)$$

$$N^g = (x_2^g - x_1^g) \times (y_2^g - y_1^g). \quad (6)$$

预测框与真实框的交集  $I$  为:

$$I = \begin{cases} (x_2^l - x_1^l) \times (y_2^l - y_1^l), & \text{if } (x_2^l > x_1^l, y_2^l > y_1^l) \\ 0, & \text{else} \end{cases}$$

其中:  $x_1^l = \max(x_1^p, x_1^g)$ ,  $x_2^l = \min(x_2^p, x_2^g)$ ,  $y_1^l = \max(y_1^p, y_1^g)$ ,  $y_2^l = \min(y_2^p, y_2^g)$ 。

预测框与真实框的最小封闭面  $M^c$  为:

$$M^c = (x_1^c, y_1^c, x_2^c, y_2^c), \quad (7)$$

其中:  $x_1^c = \min(x_1^p, x_1^g)$ ,  $x_2^c = \max(x_2^p, x_2^g)$ ,  $y_1^c = \min(y_1^p, y_1^g)$ ,  $y_2^c = \max(y_2^p, y_2^g)$ 。

最小封闭面面积  $N^c$  为:

$$N^c = (x_2^c - x_1^c) \times (y_2^c - y_1^c), \quad (8)$$

则  $M^p$  和  $M^g$  的 IOU 和 GIOU 公式为:

$$IOU = \frac{I}{N^p + N^g - I}, \quad (9)$$

$$GIOU = IOU - \frac{N^c - (N^p + N^g - I)}{N^c}. \quad (10)$$

由式(5)~式(10)可以看出, GIOU 能更加准

确地描述预测框和真实框的距离,对训练和检测效果更好,所以,本文在原损失函数中的位置误差项使用 GIOU 替换 IOU。

### 3 实验

本文中的整个实验在以 TensorFlow 为后端的 keras 中实现,实验环境为: Cuda 9.0 加速;硬件配置为 AMD Ryzen 5 Six-Core Processor @ 3.4 GHz、GeForce GTX 1080Ti 显卡;操作系统为: Ubuntu 18.04.1 LTS。总迭代次数为 50 020,初始学习率设置为 0.001,在 40 000 次迭代之后,学习率为 0.000 1,小批量设置为 16,细分设置为 4,权重衰减系数设定为 0.000 5,动量系数设定为 0.9。训练时,模型的输入大小每次迭代改变 10 次,这样最终的模型对不同尺寸的图像具有更好的检测效果。

#### 3.1 数据集制作

针对所研究的问题,本文使用混合数据集。首先使用自动驾驶数据集 KITTI,将原始数据集 8 个类别重新划分:将汽车,卡车和货车合并为机动车类;行人,坐着的行人和骑车的人合并为人类;最后两项电车和杂项删除。最终得到的训练样本数量为:机动车 33 250 个、行人 4 900 个。削减 KITTI 中机动车数量,与 INRIA 行人数据集、VOC2007 行人数据集混合,最终得到共 50 000 张图片的混合数据集,其中车类和人类比例接近 1:1,防止数据集中样本差距过大导致的过拟合问题,其中 70% 用于训练,20% 用于验证,10% 用于测试。混合数据集背景复杂,行人姿态多样,遮挡的程度和样本目标的大小不同,可以得到更好的训练效果和模型泛化能力。

#### 3.2 实验结果与分析

为了评估改进算法的准确率,将 KITTI、INRIA、VOC、混合数据集分别在 YOLOV3, Tiny YOLOV3 和本文改进算法中训练并验证测试,计算平均精度均值(mAP)和检测速度(FPS)。测试结果如表 3 所示。

由表 3 可以看出,本文使用混合数据集训练后的检测平均精度均值高于各单个数据集,能得到更好的训练效果和检测结果。YOLOV3 的检测准确率远高于 Tiny YOLOV3 和本文改进的算法,但其网络的模型大,参数多,计算量大,从而导

致 FPS 低,不适用于汽车的嵌入式设备对目标的实时检测。对比 Tiny YOLOV3 和本文改进的算法,虽然 FPS 下降了 2.8,检测速度稍慢,但仍高于 YOLOV3,满足实时检测要求。

表 3 各算法用不同数据集训练后的测试结果

Tab. 3 Test results of algorithms trained with different data sets

数据集	算法名称	mAP/%	FPS/(frame · s <sup>-1</sup> )
KITTI	YOLOV3	83.6	23.6
	Tiny YOLOV3	60.3	35.5
	改进算法	70.1	32.7
INRIA	YOLOV3	75.8	23.6
	Tiny YOLOV3	46.9	35.5
	改进算法	65.7	32.7
VOC	YOLOV3	81.2	23.6
	Tiny YOLOV3	58.3	35.5
	改进算法	68.1	32.7
混合数据集	YOLOV3	85.1	23.6
	Tiny YOLOV3	61.4	35.5
	改进算法	71.2	32.7

表 4 各算法在混合数据集上检测结果

Tab. 4 Detection results of algorithm on a mixed data set

算法名称	类别	AP/%	mAP/%
YOLOV3	机动车	87.3	
	高度小于 16 pxl 的行人	76.5	85.1
	高度为 16~32 pxl 的行人	85.5	
	高度大于 32 pxl 的行人	86.7	
Tiny YOLOV3	机动车	75.6	
	高度小于 16 pxl 的行人	38.6	61.4
	高度为 16~32 pxl 的行人	45.3	
	高度大于 32pxl 的行人	57.7	
改进算法	机动车	76.3	
	高度小于 16 pxl 的行人	58.3	71.2
	高度为 16~32 pxl 的行人	67.1	
	高度大于 32 pxl 的行人	72.9	

表4为各算法在混合数据集上的检测结果,本文改进的算法对机动车的检测准确率比 Tiny YOLOV3 提高 0.7%,对高度小于 16 pxl、高度为 16~32 pxl、高度大于 32 pxl 的行人检测准确率比 Tiny YOLOV3 分别提高了 19.7%、21.8%、15.2%,平均准确率为 71.2%,比 Tiny YOLOV3 算法 61.4% 的平均准确率提高了 9.8%,在保证实时检测的前提下,提高了目标检测精度。

图4是本文算法的真实场景检测结果,并与 Tiny YOLOV3 检测结果进行比较。其中图4(a)~图4(e)为 Tiny YOLOV3 检测结果,图4(f)~图4(j)为相同情况下改进算法的检测结果。a1

中,漏检了左侧高度小于 16 pxl 的行人,图4(f)中无行人漏检;图4(b)中将两个高度小于 16 pxl 的行人检测为一人,图4(g)中正确检测为两人;图4(c)中,漏检了远处高度为 16~32 pxl 的行人,图4(h)中无行人漏检;图4(d)中漏检了汽车遮挡身体的高度为 16~32 pxl 的行人,图4(i)准确检测出行人;图4(e)中,漏检了右侧高度小于 16 pxl 和高度为 16~32 pxl 的行人,图4(j)中无行人漏检。可以看出,本文改进算法对行人等小目标有更好的检测效果,并且在光线较弱的环境下仍有良好的检测结果,这表明改进算法对复杂实时环境具有良好的适应性,能更准确地检测目标。

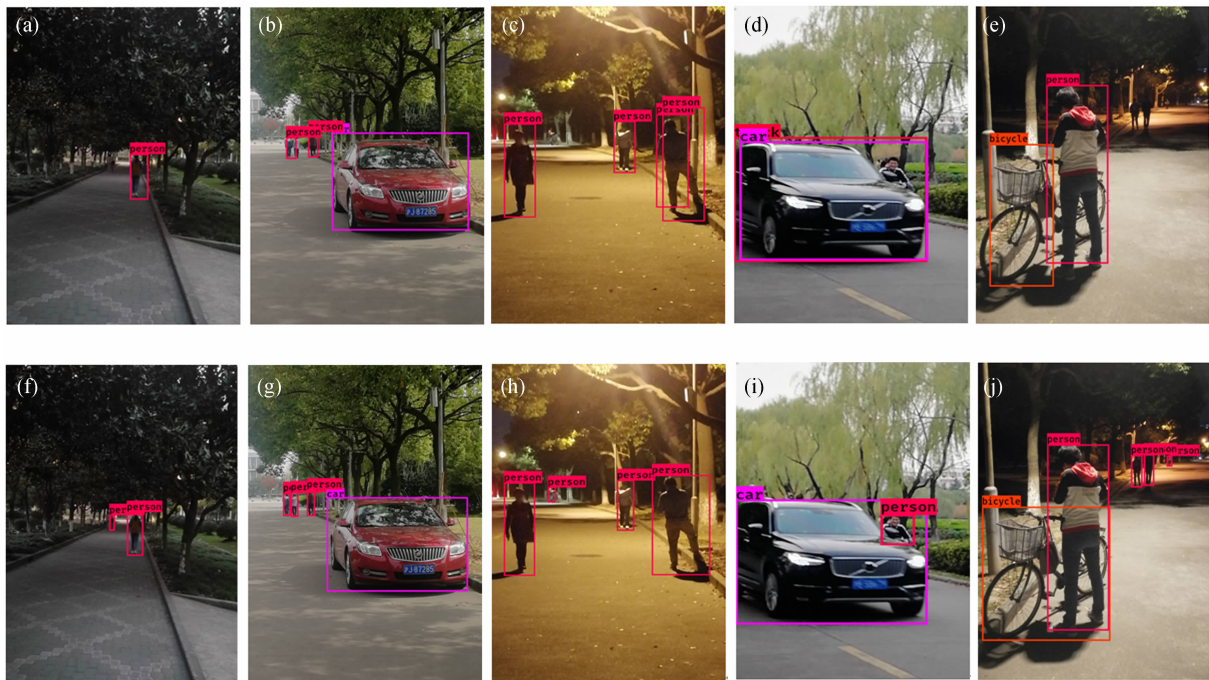


图4 实验结果

Fig. 4 Experimental results

## 4 结论

本文改进的 Tiny YOLOV3 算法,不仅通过在特征提取网络增加卷积层数提高了目标特征提取量,还通过深度可分离卷积构造反残差块替换传统卷积降低了网络模型尺寸和参数计算量,保证目标检测速度。同时,通过在

预测网络增加尺度形成三尺度预测更好的检测行人等小目标。最后,通过优化损失函数中的边界框位置误差项进一步提高了目标检测准确率。通过在制作的混合数据集中训练后的实验结果表明:改进的 Tiny YOLOV3 算法目标检测准确率为 71.2%,比原算法提高了 9.8%,FPS 为 32.7 frame/s,满足实时检测要求。虽然改进后算法的目标检测准确率有所

提高,但仍无法满足精度要求,跟大型检测网络相比还有较大差距,如何在满足实时检测的

要求下进一步提高准确率,这是下一步工作要努力的方向。

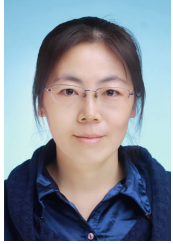
## 参考文献:

- [1] RÄTSCH G, ONODA T, MÜLLER K R. Soft margins for Ada-Boost [J]. *Machine Learning*, 2001, 42(3): 287-320.
- [2] 丁鹏, 张叶, 刘让, 等. 结合形态学和 Canny 算法的红外弱小目标检测[J]. *液晶与显示*, 2016, 31(8): 793-800.
- DING P, ZHANG Y, LIU R, *et al.*. Infrared small target detection based on adaptive Canny algorithm and morphology[J]. *Chinese Journal of Liquid Crystal and Displays*, 2016, 31(8): 793-800. (in Chinese)
- [3] 张国梁, 贾松敏, 张祥银, 等. 采用自适应变异粒子群优化 SVM 的行为识别[J]. *光学精密工程*, 2017, 25(6): 1669-1678.
- ZHANG G L, JIA S M, ZHANG X Y, *et al.*. Action recognition based on adaptive mutation particle swarm optimization for SVM[J]. *Editorial Office of Optics and Precision Engineering*, 2017, 25(6): 1669-1678. (in Chinese)
- [4] 梁华, 宋玉龙, 钱锋, 等. 基于深度学习的航空对地小目标检测[J]. *液晶与显示*, 2018, 33(9): 793-800.
- LIANG H, SONG Y L, QIAN F, *et al.*. Detection of small target in aerial photography based on deep learning[J]. *Chinese Journal of Liquid Crystal and Displays*, 2018, 33(9): 793-800. (in Chinese)
- [5] GIRSHICK R. Fast R-CNN[C]. *Proceedings of the IEEE international conference on computer vision*. 2015: 1440-1448.
- [6] REN S, HE K, GIRSHICK R, *et al.*. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[C]. *Proceedings of International Conference on Neural Information Processing Systems*, 2015: 91-99.
- [7] HE K, GKIOXARI G, DOLLÁR P, *et al.*. Mask R-CNN[C]. *Computer Vision (ICCV), 2017 IEEE International Conference on. IEEE*, 2017: 2980-2988.
- [8] FU C Y, LIU W, RANGA A, *et al.*. DSSD: Deep convolutional single shot detector[J]. *arXiv preprint arXiv: 1701.06659*, 2017.
- [9] REDMON J, DIVVALA S, GIRSHICK R, *et al.*. You Only Look Once: Unified, Real-Time Object Detection[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [10] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[J]. *arXiv preprint arXiv*, 1612.08242, 2017.
- [11] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. *arXiv preprint arXiv*, 1804.02767, 2018.
- [12] LIN T Y, DOLLAR P, GIRSHICK R, *et al.*. Feature pyramid networks for object detection[J]. *arXiv preprint arXiv*: 1612.03144, 2016.
- [13] NOH H, HONG S, HAN B. Learning Deconvolution Network for Semantic Segmentation [C]. *IEEE International Conference on Computer Vision*, 2016: 1520-1528.
- [14] 刘智, 黄江涛, 冯欣. 构建多尺度深度卷积神经网络行为识别模型[J]. *光学精密工程*, 2017, 25(3): 799-805.
- LIU ZH, HUANG J T, FENG X. Construction of multi-scale deep convolutional neural network behavior recognition model [J]. *Opt. Precision Eng.*, 2017, 25(3): 799-805. (in Chinese)
- [15] 杨州, 慕晓冬, 王舒洋, 等. 基于多尺度特征融合的遥感图像场景分类[J]. *光学精密工程*, 2018, 26(12): 232-240.
- YANG YANG ZH, MU X D, WANG SH Y, *et al.*. Remote sensing image scene classification based on multiscale feature fusion[J]. *Opt. Precision Eng.*, 2018, 26(12): 232-240. (in Chinese)
- [16] 刘伟宁. 基于小波域扩散滤波的弱小目标检测[J]. *中国光学*, 2011, 4(5): 503-508.
- LIU W N. Dim target detection based on wavelet field diffusion filter[J]. *Chinese Optics*, 2011, 4(5): 503-508. (in Chinese)
- [17] 刘希佳, 陈宇, 王文生, 等. 小目标识别的小波阈值去噪方法[J]. *中国光学*, 2012, 5(3): 248-256.

LIU X J, CHEN Y, WANG W SH, *et al.*. Denoising algorithm of wavelet threshold for small

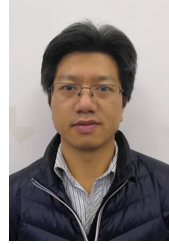
target detection[J]. *Chinese Optics*, 2012, 5(3): 248-256. (in Chinese)

#### 导师简介:



马立(1977—),女,黑龙江人,研究员,博士生导师,2007年于哈尔滨工业大学获得博士学位,现为上海大学机电工程与自动化学院教师,主要研究方向为微操作机器人技术。E-mail: malian@shu.edu.cn

#### 通讯作者:



欧阳航空(1976—),男,湖南人,工程师,2005年于上海大学获得硕士学位,现为上海大学机电工程与自动化学院教师,主要研究方向为精密运动控制技术。E-mail: hankow@shu.edu.cn

(本栏目编辑:秦思)