

基于YOLOv7-Tiny的交通标识检测算法研究

陈杨山¹, 张传庆¹, 赵曙光¹, 刘西钉², 王建强³

¹东华大学信息科学与技术学院, 上海

²上海琴信科技有限公司, 上海

³上汽大众汽车有限公司, 上海

收稿日期: 2023年3月15日; 录用日期: 2023年4月13日; 发布日期: 2023年4月20日

摘要

目标检测是智能驾驶系统中的重要组成部分, 其中交通标识属于小目标检测, 在图像中所占的像素比极少, 识别难度大, 且种类繁多, 数据样本不均衡, 因此本文提出一种基于YOLOv7-tiny的交通标识检测方法。在YOLOv7-tiny算法中增加目标检测层, 使网络更加关注小目标, 改善目标检测的效果。引入TAM (Triplet Attention Module)三重注意力机制, 通过使用一个三支结构计算注意力权重, 提高小目标检测能力。在TT100K交通标识数据集上, mAP达到84.23%, 相较于YOLOv7-tiny提高了4.21%。试验结果表明, 该方法对于复杂环境下的交通标识具有更好的检测性能, 更能满足实际需求。

关键词

YOLOv7-Tiny, TAM, 小目标检测层, TT100K

Research on Traffic Sign Detection Algorithm Based on YOLOv7-Tiny

Yangshan Chen¹, Chuanqing Zhang¹, Shuguang Zhao¹, Xiding Liu², Jianqiang Wang³

¹College of Information Science and Technology, Donghua University, Shanghai

²Shanghai Qinxin Technical Co., Ltd., Shanghai

³SAIC Volkswagen Automotive Co., Ltd., Shanghai

Received: Mar. 15th, 2023; accepted: Apr. 13th, 2023; published: Apr. 20th, 2023

Abstract

Object detection is an important part of intelligent driving system, among which traffic signs are small target detection, accounting for a very small pixel ratio in the image, which is difficult to iden-

文章引用: 陈杨山, 张传庆, 赵曙光, 刘西钉, 王建强. 基于 YOLOv7-Tiny 的交通标识检测算法研究[J]. 计算机科学与应用, 2023, 13(4): 737-744. DOI: 10.12677/csa.2023.134072

tify, with numerous types and unbalanced data samples. Therefore, a traffic sign detection method based on YOLOv7-tiny is proposed in this paper. Add the object detection layer to the YOLOv7-tiny algorithm to make the network pay more attention to small targets and improve the effect of target detection. The Triplet Attention Module (TAM) is introduced to improve the detection ability of small targets by using a three-branch structure to capture interdimensional interactions and calculate attention weights. In TT100K traffic sign data set, mAP reaches 84.23%, 4.21% higher than YOLOv7-tiny. The test results show that this method has better detection performance for traffic signs in a complex environment and can meet the actual demand.

Keywords

YOLOv7-Tiny, TAM, Small Object Detection Layer, TT100K

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

交通标识识别是智能驾驶感知系统中的基础功能之一，对于路径规划以及行车安全具有重要作用，帮助智能驾驶汽车在行驶阶段做出准确及时的决策。在实车行驶过程中，视觉感知易受到光照环境、天气变化等条件的影响，出现漏检误检现象，无法满足智能驾驶安全需求[1]。随着深度学习的发展，网络层越深，参数量也越多，图像识别与目标检测得到了进一步的发展。现阶段，目标检测算法主要分为两大类，基于候选区域生成的一阶段方法和基于回归的二阶段方法。二阶段算法以 Faster R-CNN 为主，检测精度高，但检测速度慢，实时性较差。一阶段算法以 YOLO 系列和 SSD 算法为代表[2]。最新的 YOLOv7-tiny 模型，参数以及运算量都较小，实时性好，且有较高的检测精度，适用于低算力平台的智能驾驶机车，满足低阶的 L2 级别智能驾驶功能，因此本文将基于 YOLOv7-tiny 轻量级目标检测算法作为框架，增加小目标检测层，加强对小目标的检测能力。同时，引入 TAM 注意力机制，增加特征提取过程中网络的跨维度信息交互能力。

2. YOLOv7-Tiny 算法介绍

YOLOv7-tiny 是 YOLOv7 [3]的轻量模型，主要面向边缘 GPU 架构，相比 YOLOv7，它的网络结构相对简单，计算量小，因此实时性出色，并且降低了对硬件的要求，更适用于移动端设备。YOLOv7-tiny 的具体网络结构如图 1 所示。

3. YOLOv7-Tiny 算法改进

3.1. 增加小目标检测层

深入分析 YOLOv7-tiny 目标检测算法，并针对应用于交通标志检测进行改进。交通标识检测效果不好的一个原因是目标样本的尺寸较小，较深的特征层无法学习到小目标的特征信息，因此提出增加小目标检测层对浅层特征图与深层特征图拼接后进行检测[4]。首先，在主干网络第 3 层就开始特征加强，获得尺寸为 160×160 的大尺寸特征图，以更好地实现小目标检测。其次，在颈部加强特征提取网络中增加一次上采样，同时增加一次特征融合，丰富小目标的图像信息。最后，针对主干网络第三层的特征提取层，增加目标检测头，在 160×160 尺寸的图像上，可以减少小目标的信息损失，提升检测效果。同时，

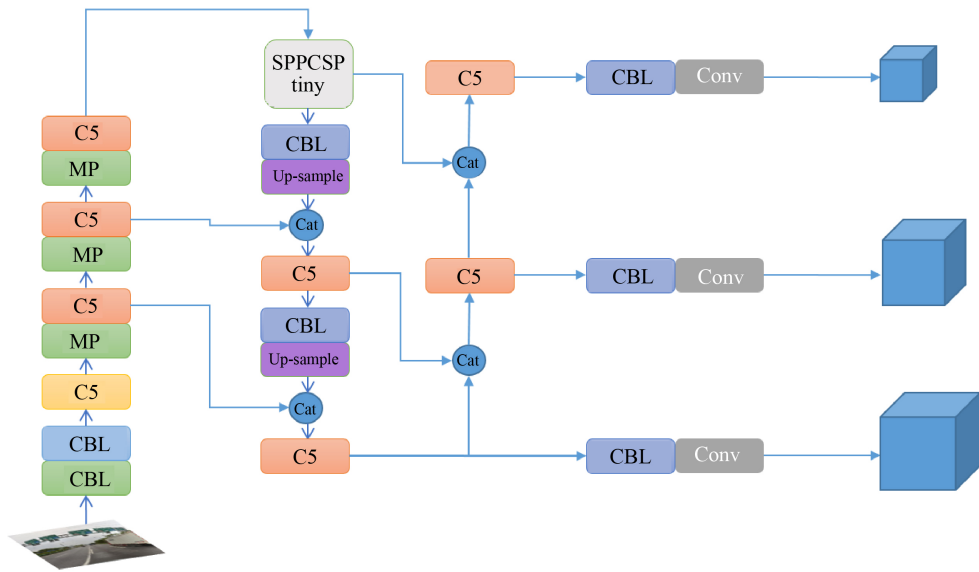


Figure 1. Diagram of YOLOv7-tiny network structure
图 1. YOLOv7-tiny 网络结构图

添加针对小目标的目标框，原小目标框参数为[10, 13, 16, 30, 33, 23]，检测头数量为三个，改进后，检测头数量为四个，小目标框参数为[5, 6, 8, 14, 15, 11]，可以让网络更加关注小目标的检测，进而提升检测效果。具体改进后的网络结构图，可参见图 4。

3.2. 引入 TAM 注意力机制

TAM (Triplet Attention Module) [5]是一种通过使用三支结构跨维度交互来计算注意力权重的方法，TAM 三重注意力机制将输入的张量进行旋转操作，然后使用残差结构建立不同维度之间的关系，同时以较小的计算量对通道信息以及空间信息进行编码。相对于通道注意力机制或空间注意力机制，TAM 三重注意力机制可以很好地进行跨通道信息的交互，从而消除了通道和权重之间的间接对应，因此应用于目标检测模型中，有显著的性能改进。具体三支结构如下图 2 所示。

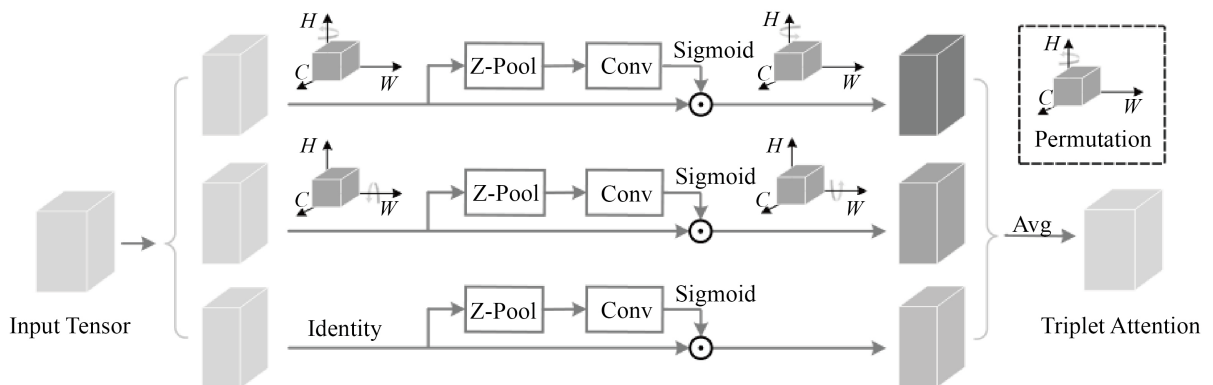


Figure 2. Illustration of three branches in triplet attention
图 2. 三重注意力的三支说明

如上图 2 所示，Triplet Attention 主要包含 3 个分支，图中最下方的分支不进行旋转操作，捕获空间

维度 H 和 W 的关系信息, 使用残差结构, 进行空间注意力权重计算, 最后使用 Sigmoid 函数进行激活操作。中间的分支, 保持空间维度 W 不变, 进行旋转操作, 捕获通道 C 维度和空间维度 H 的信息, 进行通道交互。最上方的分支, 则是捕获通道 C 维度与空间维度 W 的跨通道交互信息。最后将三支进行单元和平均操作。

Z-Pool 通过将平均池化和最大池化特征连接, 使得张量的第 0 维减少到 2 维。这使得该层能够保留实际张量的信息, 同时缩小其深度, 从而进一步地减少计算量。具体公式如下所示:

$$Z\text{-pool}(\chi) = [\text{MaxPool}_{0d}(\chi), \text{AvgPool}_{0d}(\chi)] \quad (1)$$

其中, $0d$ 是进行最大池化和平均池化操作的第 0 维。例如, 一个形状张量 $(C \times H \times W)$ 的 Z-Pool 结果是一个形状张量 $(2 \times H \times W)$ 。

具体的 TAM 网络结构如下图 3 所示。右侧第一分支不进行输入张量旋转操作, 首先输入张量经过 Z-Pool, 进行通道池化操作, 随后进行 7×7 的卷积操作, 随后进行 BatchNorm 归一化操作, 最后使用 Sigmoid 激活函数生成空间注意力权重。相对于右侧第一分支, 另外两分支首先进行 Permute 旋转操作, 再进行后续的跨通道交互操作。

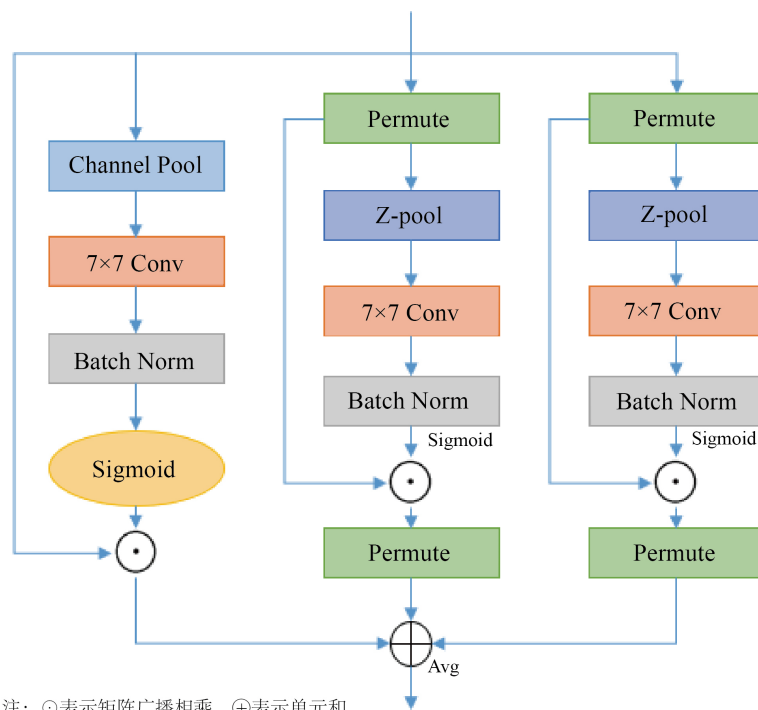


Figure 3. Network structure of TAM
图 3. TAM 网络结构

3.3. 改进后的 YOLOv7-Tiny 网络结构

在原 YOLOv7-tiny 主干特征提取网络增加一层浅层特征层, 针对增加的浅层特征层, 在加强特征提取网络中增加一次上采样操作, 并与这一特征层进行特征融合操作, 获得 160×160 的特征图。同时, 针对增加的浅层特征层, 增加了 160×160 的目标检测头。在主干特征提取网络中提取出的四个有效特征层上增加 TAM 三重注意力机制, 以提升检测精度, 具体网络结构如下图 4 所示。

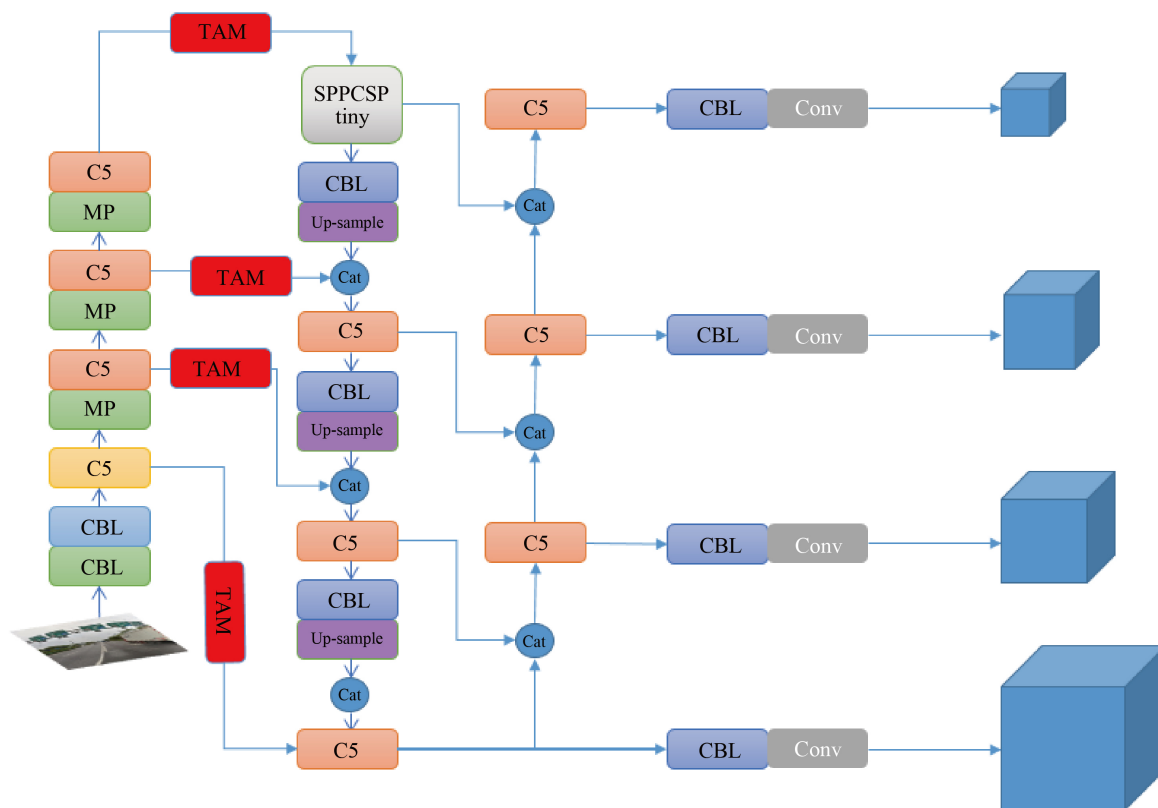


Figure 4. Diagram of YOLOv7-tiny network structure after improvement
图 4. 改进后的 YOLOv7-tiny 网络结构图

4. 实验结果及分析

4.1. TT100K 数据集

本文的实验采用交通标志数据集 Tsinghua-Tencent 100K (简称 TT100K) [6], 该数据集包含 221 个类别, 26349 个目标, 图像尺寸为 2048×2048 。该数据集存在两个问题: 首先目标尺度小的问题, 目标在图像中所占比例较小; 其次就是类别不均衡的问题, 有的类别多达两千多个, 而有的只有几个、十几个。这些都对检测和识别带来了很大挑战。这里选择 221 个类别中数量最多的 45 个类别进行实验, 并按照 6:2:2 的比例重新划分训练集验证集测试集[7]。

4.2. 实验环境及参数设置

实验环境: 操作系统 Ubuntu 20.04, 深度学习框架 Pytorch 1.10.0 + Cuda 11.3 + cudnn 8200, CPU 为 AMD EPYC 7642, 内存 80 GB; GPU 为单块 NVIDIA GeForce RTX 3090, 显存为 24 GB。

参数设置: 图片输入大小 640×640 , 训练时采用 Adam 优化器, 初始学习率 0.001, 采用余弦退火衰减策略调整学习率, 动量大小 0.937, 权重衰减系数为 0.0005, 批处理尺寸(Batchsize)设为 64, 迭代 300 个 epoch。

4.2.1. 实验结果

改进后的 YOLOv7-tiny 模型训练结果如下, 准确率(Precision)为 0.831, 置信度为 0 时的召回率(Recall)为 0.802, 平均精度值(mAP)为 0.842, 训练结果如下图 5 所示。

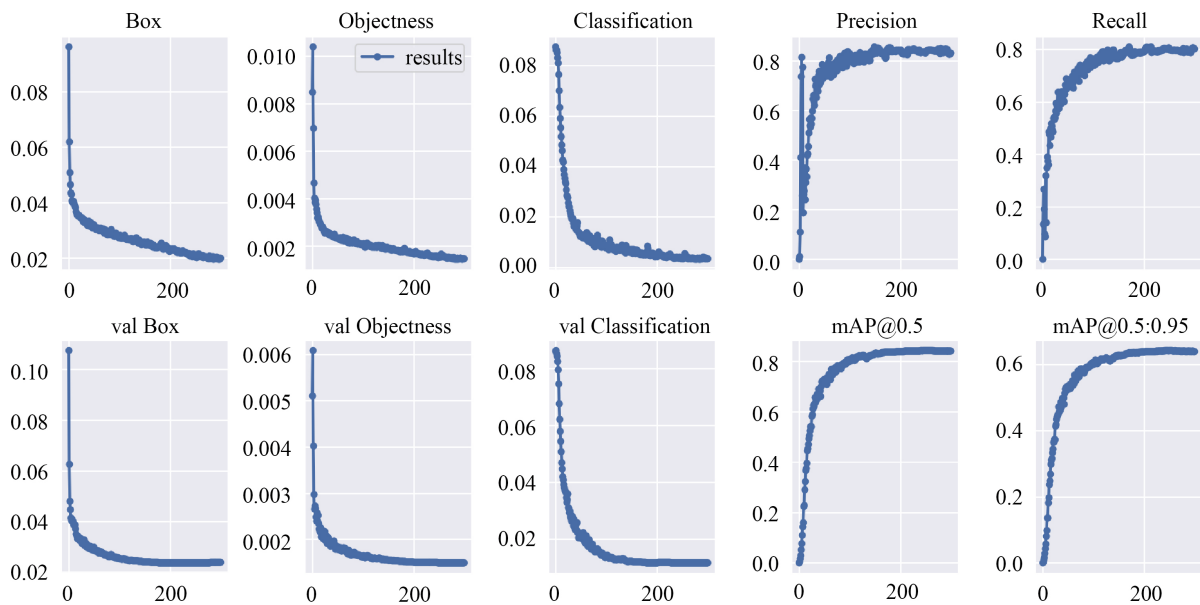


Figure 5. Results of training
图 5. 训练结果

由于原图像尺寸为 2048×2048 ，图幅较大，交通标识所占尺寸较小，为便于更加直观对比，截取原始图片中交通标识部分进行对比分析。由于复杂路口交通标识密集，目标框重叠，因此在模型进行检测过程中，置信度值设置为 0.5，因此置信度值低于 0.5 的目标框将不再显示。图 6 展示了模型在复杂路口检测结果，图 7 展示了在高速路段的检测结果，详细对比可参照下方图片。从对比图片中可清晰看出，改进后的模型检测精度有显著提升。

4.2.2. 实验对比

为验证算法模型的可靠性以及准确性，本文采用平均精度值 mAP 作为评价标准。平均精度值 mAP 为模型对各类别检测精度 AP 的平均值，代表模型的整体检测精度[8]。首先与相近参数量及计算量的 YOLOv5s 进行对比分析。由表 1 可见，YOLOv7-tiny 相对于 YOLOv5s 对于交通标识的检测平均精度提升了接近 2%。改进后的算法模型，相对于 YOLOv5s 平均检测精度提升了 6.13%，相对于 YOLOv7-tiny 提升了 4.21%，对于交通标识这类小目标有显著改进。

根据表 1 中数据对比，以及图 8 的结果展示，YOLOv7-tiny 轻量级目标检测模型以及改进后的模型显著提升了交通标识类的小目标检测精确度。



Figure 6. Comparison of detection results at complex intersection
图 6. 复杂路口检测结果对比



Figure 7. Comparison of detection results in highway
图 7. 高速路段检测结果对比

Table 1. Comparison experiment
表 1. 对比实验

| 模型 | mAP |
|-------------|--------|
| YOLOv5s | 78.13% |
| YOLOv7-tiny | 80.02% |
| Ours | 84.23% |

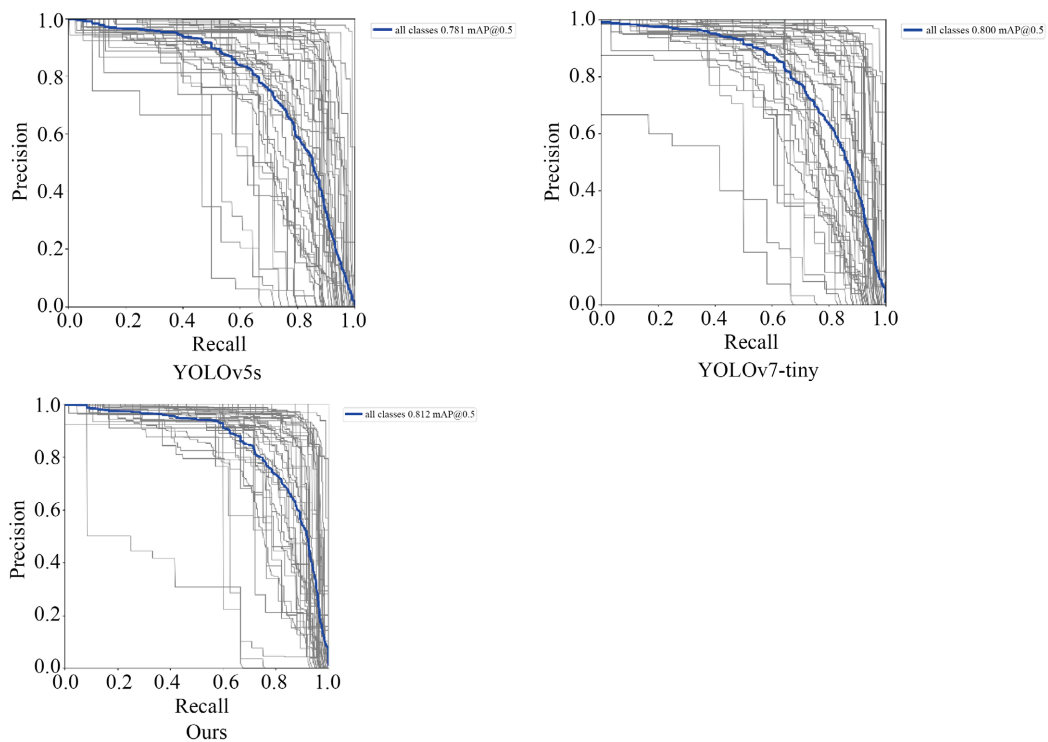


Figure 8. Comparison of results from different models
图 8. 不同模型结果对比

针对 TAM 三重注意力机制以及增加小目标检测层对于模型的影响, 通过消融实验验证, 两者都可以提高模型的检测精度, 具体检测精度值可参见下表 2, 将两者同时引入 YOLOv7-tiny 模型中, 可以显著提高模型的检测精度。

Table 2. Ablation experiment
表 2. 消融实验

| 模型 | mAP |
|-------------------------|--------|
| YOLOv7-tiny | 80.02% |
| YOLOv7-tiny + TAM | 81.21% |
| YOLOv7-tiny + 检测层 | 82.53% |
| YOLOv7-tiny + TAM + 检测层 | 84.23% |

5. 结论

本文针对轻量级目标检测网络 YOLOv7-tiny 改进, 在原主干特征提取网络增加浅层特征层, 以提高小目标的检测能力, 同时针对这一浅层特征层增加一次上采样操作, 并进行特征融合操作, 丰富小目标的信息, 同时增加目标检测头, 以及对应的小目标检测框, 提高小目标的特征提取能力。增加特征提取层之后, 在主干网络获得的四个有效特征提取层中引入 TAM 三重注意力机制, 利用其三支结构, 实现跨维度的权重信息交互, 进而提高小目标检测的精确度。经过试验验证分析, 相对于经典网络模型 YOLOv5s, 改进后的平均精度提高了 6.13%, 相对于原模型, 平均精度值上升了 4.21%。改进后的模型延续了 YOLOv7-tiny 出色的实时性特性, 同时具有更出色的小目标检测效果, 适用于低计算能力的移动设备。

参考文献

- [1] 郭朦, 陈紫强, 邓鑫, 等. 基于 YOLOv5l 和 ViT 的交通标志检测识别方法[J]. 科学技术与工程, 2022, 22(27): 12038-12044.
- [2] 刘非. 基于 YOLOv4-Tiny 的交通标志检测[J]. 信息技术与信息化, 2021(5): 18-20.
- [3] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2022) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. ArXiv: 2207.02696.
- [4] 王鹏飞. 复杂道路背景下的目标检测算法研究[D]: [硕士学位论文]. 桂林: 广西师范大学, 2022.
- [5] Misra, D., Nalamada, T., Arasanipalai, A.U. and Hou, Q.B. (2021) Rotate to Attend: Convolutional Triplet Attention Module. 2021 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, 3-8 January 2021, 3139-3148. <https://doi.org/10.1109/WACV48630.2021.00318>
- [6] Zhu, Z., Liang, D., Zhang, S., et al. (2016) Traffic-Sign Detection and Classification in the Wild. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 27-30 June 2016, 2110-2118. <https://doi.org/10.1109/CVPR.2016.232>
- [7] 孟繁星, 于臻. 基于 YOLOv5-EA 的交通标志识别[J]. 辽宁工业大学学报(自然科学版), 2022, 42(5): 303-310.
- [8] 梁泓, 赵曙光. 基于 Faster-RCNN 的道路异常状态检测方法研究[J]. 计算机科学与应用, 2022, 12(3): 546-553. <https://doi.org/10.12677/CSA.2022.123055>