

# 基于自适应融合的实时车辆检测

陈 婷<sup>1</sup>, 朱熟康<sup>1</sup>, 高 涛<sup>1</sup>, 李 浩<sup>2</sup>, 涂辉招<sup>2</sup>, 李子琦<sup>1</sup>

(1. 长安大学 信息工程学院, 陕西 西安 710064; 2. 同济大学 交通运输工程学院, 上海 201804)

**摘要:** 针对传统的车辆检测技术检测速度慢和精度低的问题, 提出了一种融合注意力的自适应金字塔网络的交通目标检测算法(fusion attention adaptive pyramid network, FAAP-Net), 可以显著降低交通事故的发生率。为了降低计算复杂度, 设计了一种轻量级的互补池化结构(CPS), 该结构在宽度和高度上采用了两组不同的池化组合, 在保持高精度的同时, 显著降低了网络的浮点运算数(GFLOPs)和参数量。为了解决智能交通系统特征图生成过程中的信息损失问题, 通过将自适应注意力模块(AAM)和特征增强模块(FEM)引入自适应融合特征金字塔网络(AF-FPN), 以融入车辆检测的形状特征。针对车辆细节特征表征弱的问题, 引入了一种按通道维度分组的注意力(SA)机制, 以增强主干网络对不同车辆检测细节特征的关注, 有效提取车辆细节的显著特征。在BDD100K数据集上的实验结果表明, FAAP-Net算法相比于传统算法, 平均精度从30.3%提升到43.7%。

**关键词:** 目标检测; 车辆检测; 互补池化; 自适应融合; 通道维度分组注意力

中图分类号: TP391.4

文献标志码: A

## Real-time Vehicle Detection Based on Adaptive Fusion

CHEN Ting<sup>1</sup>, ZHU Shukang<sup>1</sup>, GAO Tao<sup>1</sup>, LI Hao<sup>2</sup>,  
TU Huizhao<sup>2</sup>, LI Ziqi<sup>1</sup>

(1. School of Information Engineering, Chang'an University, Xi'an, Shaanxi 710064, China; 2 College of Transportation Engineering, Tongji University, Shanghai 201804, China)

**Abstract:** A traffic target detection algorithm, fusion attention adaptive pyramid network (FAAP-Net), is proposed to address the issues of slow speed and low accuracy in traditional vehicle detection techniques, significantly reducing the occurrence of traffic accidents. To mitigate computational complexity, a lightweight

complementary pooling structure (CPS) is designed, employing two sets of different pooling combinations in width and height, which maintains a high precision while significantly reducing the floating point operations per second (GFLOPs) and the parameter count of the network. Addressing the information loss during intelligent traffic system feature map generation, the adaptive fusion feature pyramid network (AF-FPN) incorporates the adaptive attention module (AAM) and the feature enhancement module (FEM) to integrate shape features for vehicle detection. Lastly, to address the weak representation of vehicle detail features, a channel-wise grouped attention (SA) mechanism is introduced, enhancing the focus of the backbone network on various vehicle detection details and effectively extracting significant features. The experimental results on the BDD100K dataset demonstrate that the FAAP-Net algorithm achieves a notable improvement, increasing the average precision from 30.3 % to 43.7 %.

**Keywords:** object detection; vehicle detection; complementary pooling; adaptive fusion; shuffle attention

在智能交通系统的发展中, 车辆检测<sup>[1-2]</sup>作为关键技术之一, 对于实现道路安全和提高交通效率具有重要意义。然而, 传统的车辆检测技术由于检测速度慢和精度低<sup>[3-5]</sup>等问题, 难以满足实时性和准确性的要求。在国内外的最新相关研究动态中, 许多学者们纷纷尝试改进车辆检测算法, 以提高其性能。但自动驾驶路测在复杂交通环境、不良横纵道路线形条件及逆光环境下的适应性较差<sup>[6]</sup>。运行环境复

收稿日期: 2023-08-28

基金项目: 国家重点研发计划(2023YFB2504703, 2019YFE0108300); 国家自然科学基金(52172379, 62001058); 中央高校基本科研业务费专项资金(300102241201, 310833160212)

第一作者: 陈 婷, 教授, 博士生导师, 工学博士, 主要研究方向为智能交通系统, 模式识别。

E-mail: tchenhd@126.com

通信作者: 李 浩, 教授, 博士生导师, 工学博士, 主要研究方向为交通风险评估, 智能网联汽车, 动态交通路网建模。

E-mail: haolij@tongji.edu.cn



论文  
拓展  
介绍

杂度也显著影响了自动驾驶行车风险<sup>[7]</sup>。近几年围绕交通检测,开展了一系列研究<sup>[8-9]</sup>,大多数方法在速度和精度之间<sup>[10-11]</sup>需要做出权衡。

为了克服传统车辆检测技术的局限性,实现高效率和高精度的车辆检测,提出了一种融合注意力的自适应金字塔网络的交通目标检测算法(fusion attention adaptive pyramid network, FAAP-Net)。

首先,引入了一种轻量级的互补池化结构(CPS),这个结构的独特之处在于它在宽度和高度两个维度上采用了两组不同的池化组合,不仅提高了检测精度,还显著减少了网络所需的参数量。其次,为了解决智能交通系统特征图生成过程中信息损失的问题,引入了一种新的方法,这个方法包括将自适应注意力模块(AAM)和特征增强模块(FEM)纳入自适应融合特征金字塔网络(AF-FPN)中,以融合车辆检测的形状特征。最后,为了进一步增强对小型交通目标的检测性能,引入了一种按通道维度分组的注意力机制(SA)的模块,这个模块有助于加强主干网络对各种车辆检测细节特征的关注,从而有效地提取出显著的车辆细节。实验结果表明,提出的FAAP-Net算法能够有效解决复杂道路环境下由于交通小目标密集度高、交通目标与背景区分度低导致的漏检、误检和遮挡严重等问题。

## 1 相关工作

交通车辆检测是智能交通系统和自动驾驶领域的重要研究方向。早期的交通车辆检测方法主要依赖于传统的计算机视觉技术,如边缘检测、色彩分割和特征提取<sup>[12-13]</sup>等。然而,这些方法在复杂的交通场景和变化多样的车辆类型下存在局限性。

随着机器学习技术的发展,研究者开始将其应用于交通车辆检测中。其中,基于支持向量机(SVM)和决策树<sup>[14]</sup>等传统机器学习算法的方法被广泛研究。这些方法通过学习大量车辆和非车辆样本的特征,实现对车辆的自动检测。近年来,深度学习技术的快速发展给交通车辆检测带来了革命性的影响。特别是卷积神经网络(CNN)在目标检测领域的应用取得了显著成果。研究者提出了多种基于深度学习的目标检测算法,如Faster R-CNN、YOLO和SSD<sup>[15]</sup>等,这些算法在准确性和实时性方面取得了显著的进步。

除了传统的图像方法,还有一些研究关注于利用立体视觉和点云数据<sup>[16]</sup>进行车辆检测。通过获取

道路场景的三维信息,可以更加准确地识别车辆和估计其运动状态。为了进一步提高车辆检测的性能和鲁棒性,一些研究探索了多传感器融合的方法。通过结合图像、雷达、激光雷达<sup>[17-19]</sup>等多种传感器数据,实现更全面和可靠的车辆检测。

轻量化车辆检测<sup>[20]</sup>是近年来在计算机视觉领域备受关注的研究方向。随着智能交通系统和自动驾驶技术的发展,对于实时性和高效性的要求越来越高,因此轻量化车辆检测成为解决这一问题的重要途径。

传统的车辆检测方法<sup>[21]</sup>通常依赖于复杂的模型和大量的参数,导致计算复杂度高,运行速度慢,并且不适用于资源有限的嵌入式设备。而轻量化车辆检测旨在通过减少模型复杂度和参数量,提高检测速度和实时性,以满足实际应用的需求。目前,学者们提出了各种创新的方法来实现轻量化车辆检测,如基于深度学习的网络设计、模型剪枝和量化、特征选择和融合<sup>[22]</sup>等技术。这些方法在提高检测速度的同时,也在一定程度上保持了检测精度,取得了显著的成果。尽管目前已有许多方法取得了良好的效果,但在实际应用中,轻量化车辆检测仍然面临着一些问题,如何在速度和精度之间进行权衡,如何适应不同场景和复杂环境<sup>[23]</sup>等。

## 2 FAAP-Net 网络结构

### 2.1 FAAP-Net 交通目标检测算法网络模型

YOLO<sup>[24]</sup>的核心思想是将目标检测问题转化为一个回归问题,模型只需要单一的前向传播,即可同时完成检测和分类。虽然YOLOv7算法的检测速度和检测精度较好,但在密集目标区域上的定位精度较低<sup>[25]</sup>。为此,提出了基于自适应融合特征金字塔的车辆检测网络FAAP-Net,对主干网络采用改进压缩模型<sup>[26]</sup>、对特征金字塔网络采用自适应注意力模块和特征增强模块<sup>[27]</sup>,在主干网络与颈部之间引入了一种按通道维度分组的注意力机制。

FAAP-Net首先对输入的BDD100K图片的尺寸统一缩放为 $640 \times 640$ 大小,再输入到互补池化结构(CPS)的轻量化主干网络中。同时在特征提取网络使用通道注意力(SA)和自适应注意力模块(AAM)及特征增强模块(FEM),在主干部分获得的3个有效特征层会在下一步加强特征提取网络部分进行特征融合,目的是融合不同图片尺度的特征信息。然后经预测网络输出3层不同尺寸大小的特征

层,输出预测结果。以BDD100K为例,输出10个不同类别,每个输出 $(x, y, w, h, o)$ 包括坐标位置和前后背景,共3个anchor数量,因此每一层的输出为 $(10+5) \times 3 = 45$ 再乘上特征层的大小即为最终的输出。

## 2.2 结合互补池化结构(CPS)的FAAP-Net

针对网络的计算复杂度和强调显著的车辆特征同时保车辆特征的细节问题,使用了一种新的互补池化结构CPS,它结合了CA<sup>[28]</sup>和MAXCA(在 $W$ 和 $H$ 方向上应用最大池化)的优点。CPS-Net作为检测器的骨干特征提取网络,采用一组轻量级的互补池化结构CPS,取代CA中的原始全局平均池化。第一组在 $W$ 方向上使用最大池化,在 $H$ 方向上使用平均池化。第二组在 $W$ 方向上使用平均池化,在 $H$ 方向上使用最大池化。这两种组合形成了CPS结构。

当相邻行和列的平均池化结果相似时,显著的车辆特征与周围较不显著的车辆特征接收到相同水平的关注,导致无法充分提取显著的车辆特征。通过应用MAXCA,可在 $W$ 和 $H$ 方向上进行最大池化操作,但网络无法获得全局的感知视角。虽然显著的车辆特征可以获得更高水平的关注,但由于池化结果仅由行和列的最大值确定,忽视了其他细节的车辆特征,因此网络容易出现不准确的定位问题,而CPS结构刚好可以解决上述的问题。

## 2.3 结合自适应融合(AF-FPN)的FAAP-Net

为了解决智能交通系统特征图生成过程中的信息损失问题,引入了一种改进的特征金字塔模型AF-FPN<sup>[29]</sup>。它在传统的特征金字塔网络基础上,引入了自适应注意模块<sup>[30]</sup>(AAM)和特征增强模块(FEM)。前者通过减少特征通道<sup>[31]</sup>来减少高级特征图中的上下文线索信息损失,后者增强了特征金字塔的表示能力并加快了推断速度<sup>[32]</sup>。

## 2.4 结合通道注意力(SA)的FAAP-Net

注意力机制如今已经被广泛用于卷积神经网络中,大大提升了很多任务上的性能表现。目前视觉中的注意力机制主要有两种,分别是通道注意力和空间注意力。针对车辆细节特征表征弱的问题,使用一种更加轻量但是更加高效的shuffle attention(SA)<sup>[33]</sup>模块,它也是将输入按照通道进行分组,对每组子特征,使用shuffle unit同时构建通道注意力和空间注意力。对每个注意力模块,设计了针对每个位置的注意力掩码来抑制噪声,加强有效的语义

信息<sup>[34]</sup>。这种方式主要设计了一个轻量但是有效的注意力模块SA,它将输入特征图按照通道分组并对每个分组用shuffle unit实现空间和通道注意力。

# 3 实验

## 3.1 实验条件与方法

BDD100K是一个大规模多样化的自动驾驶和计算机视觉研究用数据集。它包含了10 s时长的100 000个驾驶视频,涵盖了不同天气条件和一天中的不同时间段的各种驾驶场景。BDD100K在计算机视觉领域被广泛应用于目标检测、实例分割和语义分割等任务。采用BDD100K交通车辆作为网络训练和测试的数据集,得到的包含车辆类型的70 000张车辆图像数据集按9:1的比例划分为训练集+验证集和测试集。训练集进一步按9:1比例划分得到56 700张训练图像和6 300张验证图像,测试集包含7 000张图像。数据集中每种车辆类型的实例数量如下:汽车(642 658个)、公交车(10 489个)、人(82 342个)、自行车(6 453个)、卡车(27 030个)、摩托车(2 726个)、火车(115个)、骑手(4 089个)、标志(215 940个)和交通灯(167 727个)。

实验相关配置如表1所示<sup>[35]</sup>。其中实验相关的超参数设置如下:训练过程包括300次迭代,默认先冻结主干训练,后解冻训练,模型的最大学习率设置为 $1.0 \times 10^{-3}$ ,最小学习率为 $1.0 \times 10^{-5}$ ,FAAP-Net算法采用了带有动量参数为0.937和衰减系数为 $5.0 \times 10^{-4}$ 的随机梯度下降(SGD)优化器。

表1 实验环境配置

Tab.1 Experiment of configuration state

实验环境	详细信息
操作系统	Windows 11
CPU	Intel(R) Core(TM) i7-10700KF CPU @ 3.80 GHz
GPU	NVIDIA GeForce RTX 3070
显存	8 GB
Python版本	Python 3.8.5
深度学习框架	Pytorch 1.8.0

## 3.2 评价指标

在目标检测中,评估指标是评价目标检测算法方法好坏的重要依据。常见的评价指标包括:精确率也叫查准率 Precision、召回率又被称为查全率 Recall、平均精度 AP、均值平均精度 mAP。本实验采用以下作为评价指标,定义如下:

$$\text{Precision} = \frac{T_p}{T_p + F_p} \times 100\% \quad (1)$$



$$\text{Recall} = \frac{T_p}{T_p + F_N} \times 100\% \quad (2)$$

$$\text{AP} = \int_0^1 \text{Precision}(t) dt \quad (3)$$

$$\text{mAP} = \frac{\sum_{i=1}^n \text{AP}_i}{N} \quad (4)$$

式(1)~(4)中: $T_p$ 为真阳性,是实际为正样本的样本被分类器正确地预测为正样本的数量; $F_p$ 为假阳性,是实际为负样本的样本被分类器错误地预测为正样本的数量; $F_N$ 为假阴性,是实际为正样本的样本被分类器错误地预测为负样本的数量。

### 3.3 实验结果与分析

在实验过程中,为了评估不同模块修改和组合所实现的性能优化程度,设计了消融实验。消融实验的数据如表2所示。消融实验的数据表明,每一项改进都在不同程度上提升了最终结果。

实验2表明,引入轻量级的CPS网络导致精确度略微下降了0.8%,但显著减少了 $16.68 \times 10^6$ 的参数数量。实验3和4显示,AF-FPN和SA在精确度上都取得了显著的提升,分别使精确度提高了2.09%和1.66%。不同的组合也表现出了整体的正向优化效果。将实验7和实验1进行比较,精确度提高了4.17%,而参数量仅增加了 $0.98 \times 10^6$ 。与实验1的比较进一步突显了CPS、AF-FPN和SA组合所实现的显著的精确度提升。将实验3和实验7进行比较,精确度提高了2.08%,参数量增加了

表2 在BDD100K下不同模块对比

Tab.2 Comparison of different modules for BDD100K

实验	CPS	AF-FPN	SA	精确度/%	模型参数/ $10^6$
1	×	×	×	38.23	39.26
2	✓	×	×	37.43	22.58
3	×	✓	×	40.32	39.97
4	×	×	✓	39.89	39.53
5	✓	✓	×	41.88	23.29
6	✓	×	✓	41.29	22.85
7	×	✓	✓	42.40	40.24
8	✓	✓	✓	43.69	23.56

注:加粗数据为最优值;×表示不采用该模块;✓表示采用该模块。

$0.27 \times 10^6$ ,验证了SA作为轻量级注意力模型以最小代价提高模型准确性的能力。实验4和实验7比较,精确度提高了2.51%,模型大小仅增加了 $0.71 \times 10^6$ ,凸显了AF-FPN模块的优秀特征融合能力。最后,将实验7和实验8进行比较,CPS的作用主要体现在模型大小和速度方面,在 $16.68 \times 10^6$ 的模型参数下,仍然实现了1.29%的精确度提升。

图1为FAAP-Net训练过程的损失变化和精度图。损失变化曲线图分为训练损失图和验证损失图,网络训练初期,训练集损失下降,验证集损失下降,表明网络正在学习BDD100K数据集的车辆信息,训练一段时间后,训练集损失不变,验证集损失不变,表明网络趋向于稳定。BDD100K数据集训练过程的10个类别的精确度(mAP)为43.69%,表明FAAP-Net具有更好的检测精度。

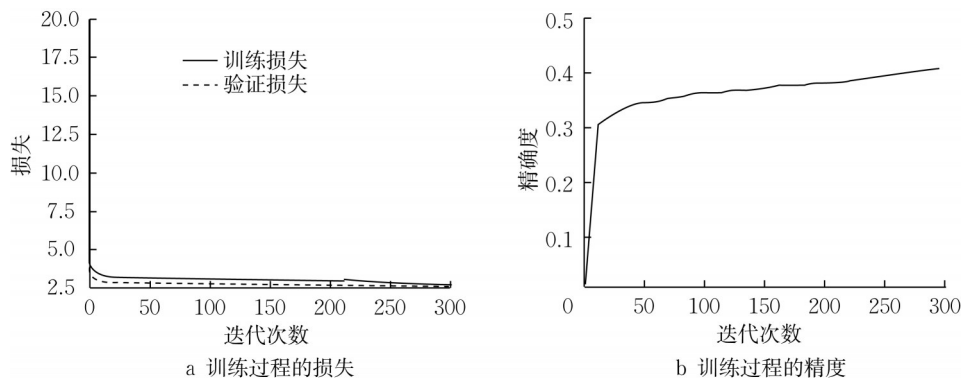


图1 训练过程的损失变化和精度

Fig.1 Loss curve and accuracy plot of training process

图2为FAAP-Net模型的真实框数量,即BDD100K数据集中每一个类别对应的目标数量,从真实值中可以观察到,测试数据集包含7 000张图像,包括10个类别:汽车、交通标志、交通灯、人、卡车、公交车、自行车、骑手、摩托车和火车。真实值还显示了每个类别的真实框数量。不难发现,

BDD100K数据集每一个类别的目标数量差异很大,即每个目标的数量变化较大。

图3为FAAP-Net的各种目标的检测精度和平均对数漏检率。AP表示网络算法模型对单个目标的检测精度,mAP表示网络模型对所有目标的平均检测精度,所有类别的mAP为43.69%,表明

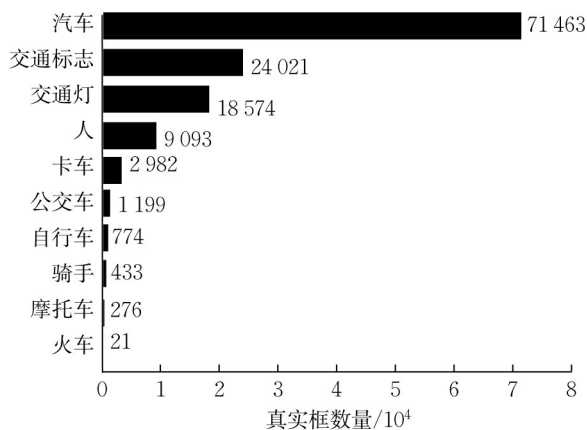


图2 训练生成的真实框

Fig.2 Training-generated real boxes

FAAP-Net整体上具有更高的平均精确度。各个类别的平均精确度也更高,表明FAAP-Net算法模型的准确性显著提高,具有较好的准确性。低平均漏检率显示,BDD100K数据集10个类别的平均对数漏检率为0.623。

图4为不同模块组合的mAP对比,可以看出通过CPS、AF-FPN和SA的有效组合能够达到43.69%的mAP,提高了车辆检测的识别准确率。

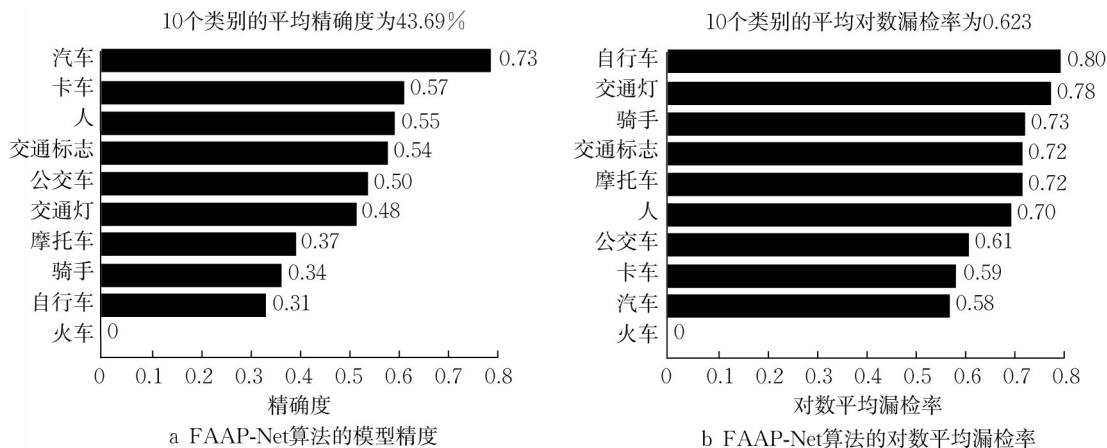


图3 FAAP-Net算法的模型精确度和对数平均漏检率

Fig.3 FAAP-Net accuracy and logarithmic average miss detection rate of improved algorithm

图5为BDD100K数据集在FAAP-Net算法上的检测结果,可以看到FAAP-Net具有更好的实时性能,FAAP-Net算法可以很容易和准确地检测到BDD100K中几乎所有的小目标。每幅图像中,都可以成功地检测到不同大小的各种车辆物体。尽管它们包含大而复杂的背景,FAAP-Net算法仍然能有效地检测到其中的物体,特别是小型车辆和密集车辆。可视化结果表明,FAAP-Net算法在BDD100K的多尺度目标检测方面具有优越的性能。

利用BDD100K数据集对SSD、Faster-RCNN、YOLOv3、YOLOv4、YOLOv5、YOLOv6、YOLOX、DETR、YOLOv7、YOLOv8以FAAP-Net算法进行了训练。结果如表3所示。从表3中可以看出,在同一BDD100K数据集上,FAAP-Net的检测精度为43.69%,参数量为 $23.56 \times 10^6$ ,GFLOPs为 $24.67 \times 10^9$ 。与SSD、Faster-RCNN、YOLOv3、YOLOv4、YOLOv5、YOLOX、YOLOv7、YOLOv8和DETR算法相比,FAAP-Net算法的准确率分别提高了5.8%、13.39%、5.46%、6.26%、3.44%、4.52%、3.71%、2.37%和3.01%。这些结果表明,本文的FAAP-Net算法具有较好的精度。在模型参数方面,YOLOv8参数尺寸最小,为 $15.16 \times 10^6$ ,FAAP-Net算法模型的参数尺寸次之,为 $23.56 \times 10^6$ ,比SSD、Faster-RCNN、YOLOv3、YOLOv4、YOLOv5、YOLOv6、YOLOX、DETR和YOLOv7模型要小得多。而且,对于帧率FPS来说,FAAP-Net算法帧率达到最高 $54.74 \text{ s}^{-1}$ ,可以做到较高的实时检测。另外,在浮点运算方面,FAAP-Net的GFLOPs为 $24.67 \times 10^9$ ,浮点运算次数仅为Faster RCNN的1/8左右。这些结果表明,FAAP-Net具有更好的实时性能和较小参数。表4展示了BDD100K数据集每一个类别不同模块组合的具体mAP精度。

为了进一步说明FAAP-Net算法对特征提取的影响,在图6中可视化了BDD100K特征图。可以看到,特征图可以捕获物体特征,将背景和噪声明显抑制,提高对小车辆检测的干扰,车辆的特征更加丰富,可以使输出特征更清晰、更准确。尽管车辆分布密集,但FAAP-Net算法显著减少了相邻车辆之间的特征混淆,有效地分离了大多数车辆的特征。这些可视化表明,FAAP-Net算法增强了车辆的特征,特别是密集车辆。具有较好的检测效果。

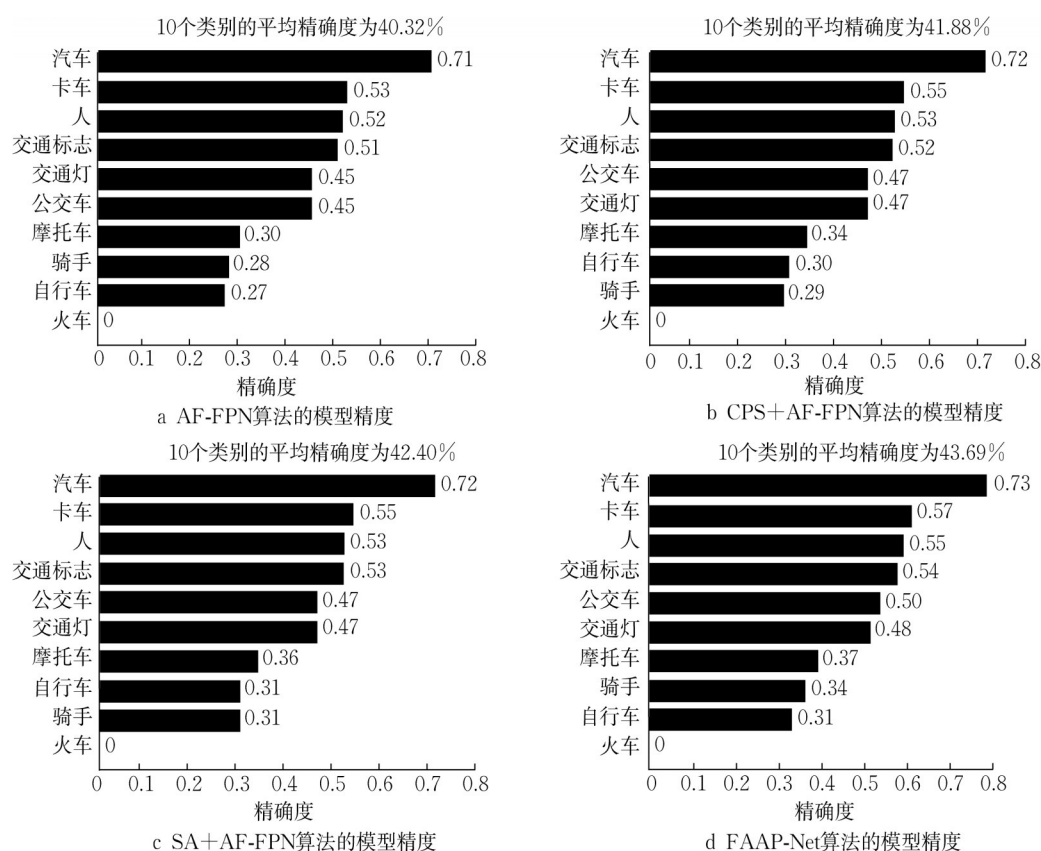


图4 不同模块组合的mAP对比

Fig.4 mAP comparison of different module combinations

表3 不同算法在BDD100K数据集上检测结果对比

Tab.3 Detection results comparison of different algorithms on BDD100K dataset

方法	精确度/%	模型参数/ $10^6$	浮点运算量/ $10^9$	帧率/ $s^{-1}$
SSD	37.89	35.87	38.47	30.53
Faster RCNN	30.30	145.16	195.32	9.67
YOLOv3	38.23	65.89	38.76	22.60
YOLOv4	37.43	68.78	35.26	24.43
YOLOv5	40.25	52.46	60.67	40.50
YOLOX	39.17	58.97	78.64	38.85
YOLOv7	39.98	39.43	55.75	39.33
YOLOv8	41.32	15.16	25.89	52.01
DETR	40.68	39.98	63.64	53.79
FAAP-Net	43.69	23.56	24.67	54.74

表4 不同算法在BDD100K数据集上各个类别的检测结果对比

Tab.4 Comparison of detection results across various classes on BDD100K dataset of different algorithms

算法	类别精确度/%										
	汽车	公交车	人	自行车	卡车	发动机	火车	骑手	标志	交通灯	平均值
SSD	60.75	51.56	39.29	34.23	50.11	28.98	0	29.58	51.05	32.35	37.89
Faster RCNN	50.12	40.58	30.52	31.56	41.48	28.25	0	27.78	48.54	4.17	30.30
YOLOv3	52.14	49.99	42.79	34.68	50.06	32.23	0	32.42	71.31	16.68	38.23
YOLOv4	51.68	45.88	39.18	33.95	50.15	30.86	0	31.65	67.36	23.59	37.43
YOLOv5	58.40	50.05	40.59	33.54	51.37	29.65	0	29.80	68.85	40.25	40.25
YOLOX	60.85	50.89	41.48	33.28	52.40	30.48	0	30.59	80.66	11.07	39.17
YOLOv7	61.32	51.02	41.21	33.47	51.01	32.21	0	31.36	79.44	18.76	39.98
YOLOv8	61.12	50.25	41.59	35.01	48.26	33.01	0	30.12	79.09	34.75	41.32
DETR	61.86	52.25	42.25	40.54	50.05	35.15	0	32.59	69.86	22.25	40.68
FAAP-Net	63.86	54.25	46.25	40.54	54.05	37.15	0	36.59	60.86	43.35	43.69





图5 BDD100K数据集在FAAP-Net算法上的检测结果

Fig. Detection results of BDD100K dataset using FAAP-Net algorithm



图6 预测结果的热力图可视化

Fig.6 Heat map visualization of prediction results

## 4 结论

提出了一种融合注意力的自适应金字塔网络的交通目标检测算法FAAP-Net。通过改进网络结构显著减少了网络所需的参数量,融合不同尺度的特征信息,提高了对交通小目标的检测能力。同时,构建自适应注意力模块(AAM)和特征增强模块(FEM)捕获融合车辆的形状通道信息和位置信息特征。此外,引入按通道维度分组的注意力机制(SA)的模块,加强主干网络对各种车辆检测细节特征的关注,从而提高提取出显著的车辆细节。

实验结果表明,所提出的FAAP-Net算法有效的提高了交通目标检测的精确性和实时性,在BDD100K数据集上,相比于YOLOv7算法,本文算法平均检测精度提升了3.71%,模型参数降低了15.87M,FAAP-Net显著提高了车辆检测的准确性和模型参数。

### 作者贡献声明:

陈 婷:总体架构,模型构建,实证分析,论文润色。  
朱熟康:数据处理,研究方法,结果讨论。  
高 涛:研究方法,思路梳理,结果分析,结论总结。  
李 浩:研究思路,指标确定,研究方法,论文润色。  
涂辉招:数据处理,平台搭建,研究方法,结果讨论。  
李子琦:数据分析,研究方法,结果讨论,论文润色。

### 参考文献:

- [1] 孔烜, 张杰, 邓露, 等. 基于机器视觉的车辆检测与参数识别研究进展[J]. 中国公路学报, 2021, 34(4): 13.  
KONG Xuan, ZHANG Jie, DENG Lu, *et al.* Research progress on vehicle detection and parameter recognition based on machine vision [J]. China Journal of Highway and Transport, 2021, 34(4): 13.
- [2] 马永杰, 马芸婷, 程时升, 等. 基于改进YOLOv3模型与Deep-SORT算法的道路车辆检测方法[J]. 交通运输工程学报, 2021, 21(2): 222.

- MA Yongjie, MA Yunting, CHENG Shisheng, *et al.* Road vehicle detection method based on improved YOLOv3 model and deep-sort algorithm [J]. *Journal of Traffic and Transportation Engineering*, 2021, 21(2): 222.
- [3] 谢光达, 李洋, 曲洪权, 等. 基于改进 Transformer 的小目标车辆精确检测算法[J]. *激光与光电子学进展*, 2022, 59(18): 364.
- XIE Guangda, LI Yang, QU Hongquan, *et al.* Precise detection algorithm for small target vehicles based on improved transformer [J]. *Laser & Optoelectronics Progress*, 2022, 59(18): 364.
- [4] 毛其超, 贾瑞生, 左羚群, 等. 基于深度学习的交通监控视频车辆检测算法[J]. *计算机应用与软件*, 2020, 37(9): 111.
- MAO Qichao, JIA Ruisheng, ZUO Lingqun, *et al.* Vehicle detection algorithm for traffic monitoring videos based on deep learning [J]. *Computer Applications and Software*, 2020, 37(9): 111.
- [5] 汪昱东, 郭继昌, 王天保. 一种改进的雾天图像行人和车辆检测算法[J]. *西安电子科技大学学报*, 2020, 47(4): 70.
- WANG Yudong, GUO Jichang, WANG Tianbao. An improved algorithm for pedestrian and vehicle detection in foggy images[J]. *Journal of Xidian University*, 2020, 47(4): 70.
- [6] 崔皓然, 涂辉招, 马杰, 等. 考虑运行环境复杂度的自动驾驶行车风险场建模方法[J]. *现代交通与冶金材料*, 2023, 3(4): 57.
- CUI Haoran, TU Huizhao, MA Jie, *et al.* Modeling method of autonomous driving risk field considering the complexity of operating environment [J]. *Modern Transportation and Metallurgical Materials*, 2023, 3(4): 57.
- [7] 涂辉招, 遇泽洋, 朱晓晖, 等. 自动驾驶路测与人工驾驶事故致因影响对比分析[J]. *现代交通与冶金材料*, 2021, 1(6): 40.
- TU Huizhao, YU Zeyang, ZHU Xiaohui, *et al.* Comparative analysis of the causative impact of autonomous driving road test and manual driving accident [J]. *Modern Transportation and Metallurgical Materials*, 2021, 1(6): 40.
- [8] 陈婷, 姚大春, 高涛, 等. 基于PreNet和YOLOv4融合的雨天交通目标检测网络[J]. *交通运输工程学报*, 2022, 22(3): 225.
- CHEN Ting, YAO Dachun, GAO Tao, *et al.* Traffic target detection network for rainy days based on prenet and YOLOv4 fusion [J]. *Journal of Traffic and Transportation Engineering*, 2022, 22(3): 225.
- [9] 高涛, 文渊博, 陈婷, 等. 基于窗口自注意力网络的单图像去雨算法[J]. *上海交通大学学报*, 2023, 57(5): 613. DOI:10.16183/j.cnki.jsjtu.2022.032.
- GAO Tao, WEN Yuanbo, CHEN Ting, *et al.* Single-image rain removal algorithm based on window self-attention network [J]. *Journal of Shanghai Jiao Tong University*, 2023, 57(5): 613. DOI:10.16183/j.cnki.jsjtu.2022.032.
- [10] 梁礼明, 熊文, 蓝智敏, 等. 改进的CornerNet Saccade车辆检测算法[J]. *重庆理工大学学报(自然科学版)*, 2021, 35(6): 137.
- LIANG Liming, XIONG Wen, LAN Zhimin, *et al.* Improved cornernet-saccade vehicle detection algorithm [J]. *Journal of Chongqing University of Technology (Natural Science)*, 2021, 35(6): 137.
- [11] 刘肯, 何姣姣, 张永平, 等. 改进YOLO的车辆检测算法[J]. *现代电子技术*, 2019, 42(13): 47.
- LIU Ken, HE Jiaojiao, ZHANG Yongping, *et al.* Improved YOLO vehicle detection algorithm [J]. *Modern Electronics Technique*, 2019, 42(13): 47.
- [12] 崔可涛, 刘怀广, 周诗洋, 等. 一种基于机器视觉的铅酸蓄电池尺寸检测方法[J]. *机床与液压*, 2021, 49(11): 97.
- CUI Ketao, LIU Huaiguang, ZHOU Shiyang, *et al.* A machine vision-based method for lead-acid battery dimension measurement [J]. *Machine Tool & Hydraulics*, 2021, 49(11): 97.
- [13] 冯润, 黄成泉, 胡雪, 等. 基于模糊拟合图像驱动的苗族服饰图像分割算法[J]. *现代纺织技术*, 2022, 30(5): 31.
- FENG Run, HUANG Chengquan, HU Xue, *et al.* Image segmentation algorithm for miao clothing based on fuzzy fitting image-driven approach [J]. *Advanced Textile Technology*, 2022, 30(5): 31.
- [14] 韩启迪, 张小桐, 申维. 基于决策树特征提取的支持向量机在岩性分类中的应用[J]. *吉林大学学报(地球科学版)*, 2019, 49(2): 611.
- HAN Qidi, ZHANG Xiaotong, SHEN Wei. Application of support vector machine based on decision tree feature extraction in lithology classification [J]. *Journal of Jilin University (Earth Science)*, 2019, 49(2): 611.
- [15] BHAVYA S B, YASHWANTH B V, NEELIMA N. An inter-comparative survey on state-of-the-art detectors—RCNN, YOLO, and SSD [C]//*Intelligent Manufacturing and Energy Sustainability: Proceedings of ICIMES 2020*. Singapore: Springer, 2021: 475-483.
- [16] 李奇明, 任杰骥, 裴晓寒, 等. 基于多模态数据协同的弱纹理表面高精度点云配准算法[J]. *光学学报*, 2022, 42(8): 117.
- LI Qiming, REN Jieji, PEI Xiaohan, *et al.* High-precision point cloud registration algorithm for weakly textured surfaces based on multi-modal data collaboration [J]. *Acta Optica Sinica*, 2022, 42(8): 117.
- [17] 龚威, 史硕, 陈博文, 等. 机载高光谱激光雷达成像技术发展与应[J]. *光学学报*, 2022, 42(12): 29.
- GONG Wei, SHI Shuo, CHEN Bowen, *et al.* Development and application of airborne hyperspectral lidar imaging technology [J]. *Acta Optica Sinica*, 2022, 42(12): 29.
- [18] 张银, 任国全, 程子阳, 等. 三维激光雷达在无人车环境感知中的应用研究[J]. *激光与光电子学进展*, 2019, 56(13): 9.
- ZHANG Yin, REN Guoquan, CHENG Ziyang, *et al.* Application research of 3d lidar in environment perception of unmanned vehicles [J]. *Laser & Optoelectronics Progress*, 2019, 56(13): 9.
- [19] 王海, 刘明亮, 蔡英凤, 等. 基于激光雷达与毫米波雷达融合的车辆目标检测算法[J]. *江苏大学学报(自然科学版)*, 2021, 42(4): 389.
- WANG Hai, LIU Mingliang, CAI Yingfeng, *et al.* Vehicle target detection algorithm based on fusion of lidar and millimeter-wave radar [J]. *Journal of Jiangsu University (Natural Science)*, 2021, 42(4): 389.



- [20] 熊李艳, 涂所成, 黄晓辉, 等. 基于 MobileVit 轻量化网络的车辆检测方法[J]. 计算机应用研究, 2022, 39(8): 2545.  
XIONG Liyan, TU Suocheng, HUANG Xiaohui, *et al.* Vehicle detection method based on Mobilevit lightweight network [J]. Application Research of Computers, 2022, 39(8): 2545.
- [21] 张炳力, 秦浩然, 江尚, 等. 基于 RetinaNet 及优化损失函数的夜间车辆检测方法[J]. 汽车工程, 2021, 43(8): 1195.  
ZHANG Bingli, QIN Haoran, JIANG Shang, *et al.* Nighttime vehicle detection method based on RetinaNet and optimized loss function[J]. Automotive Engineering, 2021, 43(8): 1195.
- [22] 申毫, 孟庆浩, 刘胤伯. 基于轻量卷积网络多层特征融合的人脸表情识别[J]. 激光与光电子学进展, 2021, 58(6): 148.  
SHEN Hao, MENG Qinghao, LIU Yinbo. Facial expression recognition based on multi-layer feature fusion of lightweight convolutional networks[J]. Laser & Optoelectronics Progress, 2021, 58(6): 148.
- [23] 刘洋, 战荫伟. 基于深度学习的小目标检测算法综述[J]. 计算机工程与应用, 2021, 57(2): 37.  
LIU Yang, ZHAN Yinwei. A review of deep learning-based small object detection algorithms [J]. Journal of Computer Engineering & Applications, 2021, 57(2): 37.
- [24] HUSSAIN M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection[J]. Machines, 2023, 11(7): 677.
- [25] 叶刚, 李毅波, 马逐曦, 等. 基于 ViBe 的端到端铝带表面缺陷检测识别方法[J]. 浙江大学学报(工学版), 2020, 54(10): 1906.  
YE Gang, LI Yibo, MA Zhuxi, *et al.* End-to-end aluminum strip surface defect detection and recognition method based on vibe[J]. Journal of Zhejiang University (Engineering Science), 2020, 54(10): 1906.
- [26] LIANG X, LIANG J, YIN T, *et al.* A lightweight method for face expression recognition based on improved MobileNetV3 [J]. IET Image Processing, 2023, 17(8): 2375.
- [27] WANG J, CHEN Y, DONG Z, *et al.* Improved YOLOv5 network for real-time multi-scale traffic sign detection [J]. Neural Computing and Applications, 2023, 35(10): 7853.
- [28] 陈加敏, 徐杨. 注意力金字塔卷积残差网络的表情识别[J]. 计算机工程与应用, 2022, 58(22): 123.  
CHEN Jiamin, XU Yang. Expression recognition based on attention pyramid convolutional residual network[J]. Journal of Computer Engineering & Applications, 2022, 58(22): 123.
- [29] 杨祥, 王华彬, 董明刚. 改进 YOLOv5 的交通标志检测算法[J]. 计算机工程与应用, 2023, 59(13): 194.  
YANG Xiang, WANG Huabin, DONG Minggang. Improved traffic sign detection algorithm based on YOLOv5 [J]. Journal of Computer Engineering & Applications, 2023, 59(13): 194.
- [30] 杨国亮, 余帅英, 杨浩. 改进 YOLOV5s 的多尺度融合口罩佩戴检测方法[J]. 计算机工程与应用, 2023, 59(14): 184.  
YANG Guoliang, YU Shuaiying, YANG Hao. Multi-scale fusion mask wearing detection method based on improved YOLOV5s [J]. Journal of Computer Engineering & Applications, 2023, 59(14): 184.
- [31] 郭列, 张团善, 孙威振, 等. 融合空间注意力机制的图像语义描述算法[J]. 激光与光电子学进展, 2021, 58(12): 329.  
GUO Lie, ZHANG Tuanshan, SUN Weizhen, *et al.* Image semantic description algorithm with integrated spatial attention mechanism [J]. Laser & Optoelectronics Progress, 2021, 58(12): 329.
- [32] 谢为俊, 魏硕, 郑招辉, 等. 采用深度多分支模型融合网络的胡萝卜缺陷识别与分割[J]. 农业工程学报, 2021, 37(2): 177.  
XIE Weijun, WEI Shuo, ZHENG Zhaohui, *et al.* Carrot defect recognition and segmentation using a deep multi-branch model fusion network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(2): 177.
- [33] ZHANG Q L, YANG Y B. Sa-net: shuffle attention for deep convolutional neural networks[C]//ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Toronto: IEEE, 2021: 2235-2239.
- [34] 周勇, 陈思霖, 赵佳琦, 等. 基于弱语义注意力的遥感图像可解释目标检测[J]. 电子学报, 2021, 49(4): 679.  
ZHOU Yong, CHEN Silin, ZHAO Jiaqi, *et al.* Explainable object detection in remote sensing images based on weak semantic attention [J]. Acta Electronica Sinica, 2021, 49(4): 679.
- [35] 王潇, 李子琦, 高涛, 等. 基于无损跨尺度特征融合的交通目标检测算法[J]. 中国公路学报, 2023, 36(9): 315.  
WANG Xiao, LI Ziqi, GAO Tao, *et al.* Traffic target detection algorithm based on non-destructive cross-scale feature fusion [J]. China Journal of Highway and Transport, 2023, 36(9): 315.