MMD-YOLOv7:黑暗条件下车辆检测方法

黄金贵,刘朋,唐文胜*

(湖南师范大学信息科学与工程学院,湖南 长沙 410081)

摘　要: 车辆检测与识别是智能交通和自动驾驶领域的一项关键技术,对于道路安全和自动驾驶起着至关重要的作用,一直是备受关注的研究热点。基于深度学习的目标检测模型,使得车辆检测精度得到大幅提高,但在夜间低光照和恶劣天气等不利条件下现有车辆检测技术的精度和可靠性仍然存在极大挑战。针对这一问题,基于YOLOv7 模型,提出一种针对夜间黑暗条件下的车辆检测方法 MMD-YOLOv7。首先基于坐标注意力(CA)机制构建一种新的多通道坐标注意力(MCCA)模块,显著提升模型在捕捉全局和局部特征信息方面的能力。其次通过构建一个多尺度卷积(MSC)模块实现对扩展高效层聚合网络(ELAN)结构的针对性改进,使得模型更好地适应夜间视觉环境中的噪声干扰,同时提升了特征提取的能力和精度。最后,引入了多分支模块(DBB),进一步增强模型对复杂特征的捕捉能力。为了验证所提出的模型效果,选取了 BDD100K 数据集中的 6 000 张夜间场景图片进行训练和测试,实验结果表明,该模型在车辆检测精度上相比原始 YOLOv7 模型提升了 5.3 百分点,展现出了模型在处理低光照情况下的强大能力。此外,在其他多个公开的车辆检测数据集上也表现出了很好的性能,验证了该模型具备很强的鲁棒性和泛化能力。

关键词: YOLO 模型;目标检测;车辆检测;深度学习;注意力机制

中图分类号: TP18　　**文献标志码:** A　　**DOI:** 10.19678/j.issn.1000-3428.0069139

MMD-YOLOv7: Vehicle Detection Method Under Dark Conditions

HUANG Jingui, LIU Peng, TANG Wensheng*

(College of Information Science and Engineering, Hunan Normal University, Changsha 410081, Hunan, China)

【Abstract】 Vehicle detection and identification is a key technology in the field of intelligent transportation and autonomous driving. It plays a vital role in road safety and autonomous driving and has always been a research hotspot. Target detection models based on deep learning have greatly improved vehicle detection accuracy, but challenges related to the accuracy and reliability of existing vehicle detection technology under adverse conditions, such as low light at night and bad weather, remain. To address these challenges, a vehicle detection method and model specifically targeted at nighttime dark conditions-MMD-YOLOv7-is proposed based on the YOLOv7 model. First, a new Multi-Channel Coordinate Attention (MCCA) module is constructed based on the Coordinate Attention (CA) mechanism, which significantly improves the model's ability to capture global and local feature information. Second, a Multi-Scale Convolution (MSC) module is innovatively designed and constructed to achieve targeted improvements to the Efficient Layer Aggregation Networks (ELAN) structure, allowing the model to better adapt to noise interference in the night visual environment, while improving the capability and accuracy of feature extraction. Finally, the Diverse Branch Block (DBB) is introduced to further enhance the model's ability to capture complex features. To verify the effectiveness of the proposed model, 6 000 night scene images in the BDD100K dataset are selected for training and testing. The experimental results show that the proposed model's vehicle detection accuracy is improved by 5.3 percentage points compared to the original YOLOv7 model, indicating its strong ability to handle low light conditions. In addition, it shows good performance on multiple public vehicle detection datasets, verifying its strong robustness and generalization capabilities.

【Key words】 YOLO model; target detection; vehicle detection; deep learning; attention mechanism

0　引言

汽车的普及使人们的生活更加便捷,但同时也带来了一系列问题,尤其是交通拥堵和出行安全等方面的问题日益严重。为了应对这些挑战,智能交通系统和自动驾驶技术的发展受到了前所未有的关注。车辆检测技术是至关重要的核心组成部分,它需要在有限的计算资源环境下,能够高度准确且快速响应地处理复杂多变的道路环境。基于深度学习的目标检测技术为实现这一目标提供了有效的解决

收稿日期: 2023-12-29　　**修回日期:** 2024-04-24

基金项目: 国家自然科学基金面上项目(62077014)。

通信作者 E-mail: * tangws@hunnu.edu.cn

计算机工程 www.ecice06.com

方案。

在实时驾驶目标检测领域，夜间环境对于检测精度提出了一系列独特且复杂的挑战。夜间的检测精度通常较低，主要归因于：1）光照条件不佳，夜间的低光照环境直接导致图像中车辆目标缺乏清晰的轮廓和细节，尤其在远处和边缘区域，这些区域的目标更容易在背景中混淆、模糊，从而使得目标检测算法难以准确判断目标的位置、形状和尺寸，此外夜间小目标如行人、标志牌等因其体积小、反射光较弱，更易于在视觉信息稀缺的环境中被遗漏，加大了检测难度；2）反光和光斑问题，夜间的反光和光斑对于多目标聚集区域尤为突出。当多个车辆聚集在一起时，它们的反光和互相照明会产生复杂的光照模式，在图像中形成干扰，使算法难以区分不同的目标；3）视野受限和光照不均匀，夜间的视野受限和光照不均匀性对远处和边缘区域的小目标检测尤其具有挑战性。不均匀的光照条件可能导致某些区域过曝，而其他区域则处于阴影中，这不仅影响目标的可见性，还增加了在复杂环境中正确识别和跟踪多个目标的难度。

然而，国内的车辆识别研究主要集中在光照条件良好的情况下，对夜间或光照不足的情况研究相对较少。针对这个问题，本文以 YOLOv7[1] 为基准模型，利用其在实时目标检测方面的良好表现，通过融合最大均值差异（MMD）方法，提出适用于黑暗状态下的车辆检测模型 MMD-YOLOv7。本文的主要贡献如下：

1）基于坐标注意力（CA）机制[2] 构建了一种新的多通道坐标注意力（MCCA）模块。通过多路径聚合的方式扩展了 CA 机制，使得模型能够更有效地捕捉全局和局部信息。

2）设计一个多尺度卷积（MSC）模块，用来替换 YOLOv7 中扩展高效层聚合网络（ELAN）结构内部的某些 CBS 模块，使得网络能更好地适应夜间噪声，特别是夜间图像中噪声较多的情况，并对 ELAN 结构进行改进优化，增强模型的特征提取能力。

3）引入文献[3]的多分支模块（DBB），采用多分支拓扑结构和多种方式的卷积增强，不仅提升了网络的特征提取能力，还有助于捕捉更复杂的特征，能够丰富特征空间。

4）为了专注于研究夜间环境下的车辆检测和识别问题，从 BDD100K 数据集[4] 中选取了 6 000 张夜间场景图片，将本文模型 MMD-YOLOv7 训练成为极端光照条件下的车辆检测模型。同时，还利用多个公开的车辆检测数据集测试本文模型性能。

1 相关工作

早期的车辆检测主要依赖于手工设计的特征和传统的机器学习算法。这些方法在复杂环境下的泛化能力和准确性非常有限。随着深度学习技术的兴起，卷积神经网络（CNN）开始在车辆检测任务中展现出强大的能力，极大地推动了这一领域的发展。

1.1 基于深度学习的车辆目标检测方法

目标检测的深度学习算法主要包括两大类：两阶段算法和单阶段算法，两种算法的主要区别在于处理候选框的方式和目标检测的流程。

两阶段算法分为两个阶段，即先提取候选框，再通过修正和分类的方式完成目标检测，从而获得较高的检测精度。主要代表是 2014 年提出的 R-CNN[5] 模型，标志着深度学习在车辆检测领域的初步应用，通过区域提议网络和深层 CNN 的结合，显著提高了检测的准确性。此后，Fast R-CNN[6]、Faster R-CNN[7]、Mask R-CNN[8] 等算法通过优化计算流程和引入更高效的区域提议机制，进一步提升了检测速度和精度。

两阶段算法由于涉及多个阶段和复杂的计算过程，其检测速度相对较慢。单阶段算法（如 YOLO、EfficientDet[9] 等）是在一次特征提取过程中直接完成目标检测，因此具有更快的检测速度，但其检测精度可能略低于两阶段算法。然而，随着技术的不断发展，单阶段算法在准确性和召回率上逐渐提升，目前在目标检测领域上单阶段算法已成为主流算法。

SSD[10]（Single Shot MultiBox Detector）和 RetinaNet[11] 等算法为解决车辆检测中的不同挑战做出了贡献，其中 SSD 通过使用不同尺度的特征图进行检测，有效地提升了对小目标车辆的检测能力，RetinaNet 则通过引入 Focal Loss 解决了前景-背景类别不平衡的问题，显著提高了检测的准确性。

YOLO 系列算法的出现为车辆检测带来了新的突破。YOLO 将车辆检测任务转换为单个回归问题，实现了高效的端到端车辆检测，其高速度和良好的准确性使其成为实时车辆检测应用的首选。随着 YOLOv2[12]、YOLOv3[13] 到最新的 YOLOv7、YOLOv8 的迭代更新，模型的性能不断优化和提高。

近年来，随着注意力机制和神经网络架构搜索（NAS）技术的发展，车辆检测算法又朝着更高效、更准确的方向迈进了一步。基于 Transformer 的模

计算机工程 www.ecice06.com

型，如 DETR[14]和其后续工作，通过直接利用自注意力机制处理检测任务，展现出对复杂场景的强大处理能力。同时，通过结合 NAS 优化网络结构，能够设计出更适合特定任务需求的高效模型。

在车辆检测领域，国外学者做了许多研究，文献[15]探讨了雨天环境对车辆检测技术的影响，并提出了有效的对策，但这些对策还缺乏具体的解决方案和模型。文献[16]介绍了一种基于雷达数据的实时区域建议算法，这不仅为车辆检测提供了新的数据维度，还提升了实时性能，但是这种方法对雷达设备的依赖性较高。文献[17]聚焦于遥感数据的研究，拓宽了数据来源和应用场景的边界，但在数据处理和分析方面较为复杂。文献[18]对 YOLOv5 进行改进，显著提升了自动驾驶领域中对小目标的检测性能，但是这项改进没有涉及交通标志的研究，且对不同应用场景可能需要进行调整和优化。

国内学者在车辆检测领域也有大量研究，主要集中在相关技术的应用和优化上。文献[19]利用深度强化学习技术开发出了一套专门针对无人驾驶车辆的避障决策框架，这一框架能够高效处理障碍物目标，但其实时性和对环境变化的适应性有待提升。文献[20]通过对模型的改进，既提升了遮挡目标的检测准确率，同时也保持了较高的检测速度，但对环境条件要求较高。文献[21]采用轻量化网络结合精细的注意力机制，有效提升了交通标志的检测效率，但在复杂的交通环境下仍存在漏检的现象。在无人机航拍领域的研究中，文献[22]通过对 YOLOv5 的改进，显著提升了目标检测的性能。

1.2 YOLOv7 模型

YOLOv7 是由 WANG 等在 2022 年开发的 YOLO 系列的模型。其检测思想和 YOLOv5 类似，采用了扩展高效层聚合网络(E-ELAN)、基于级联模型的模型缩放和模型重新参数化等策略，以实现检测效率和精度之间的良好平衡。如图 1 所示，YOLOv7 模型包括输入、主干、头部等 3 个组成部分。

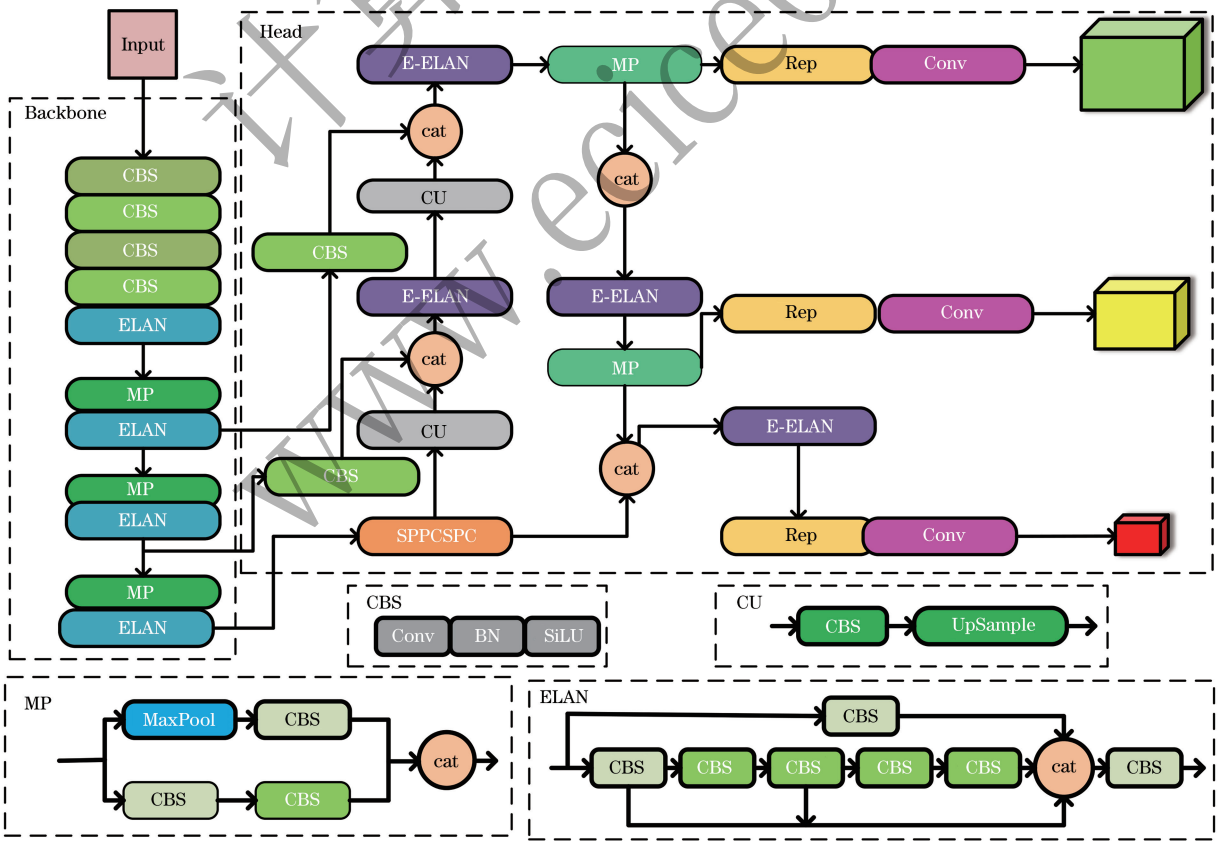


图 1 YOLOv7 模型的结构

Fig. 1 Structure of the YOLOv7 model

1)输入部分，YOLOv7 模型的预处理阶段采用了数据增强技术，并利用 YOLOv5 建立的自适应锚框计算方法，将输入图像尺寸统一放缩到 640×640 像素，以满足骨干网络输入大小的要求。

2)主干部分包括 3 个主要组件：CBS 模块，ELAN 模块和 MP 模块。CBS 模块由卷积、批量归一化和 SiLU 激活函数组成。ELAN 模块通过不同的特征组学习特征来增强网络的学习能力。MP 由 CBS 和 MaxPool 组成，MaxPool 提取小局部区域的最大值信息，而 CBS 提取小局部区域的所有值信

计算机工程 www.ecice06.com

息，从而提高网络的特征提取能力。

3）头部 Head 网络采用特征金字塔网络（FPN）架构[23]，包含跨阶段部分连接的空间金字塔池化（SPPCSPC）结构、扩展高效层聚合网络（E-ELAN）和 MP 模块。SPPCSPC 结构通过在空间金字塔池化（SPP）结构中结合跨阶段部分连接（CSPC）技术，能够对输入的特征图进行多尺度的空间金字塔池化，从而提高模型的感受野和特征表达能力。E-ELAN 在原始 ELAN 的基础上，改变计算块的同时保持原 ELAN 的过渡层结构，能够在不破坏原有梯度路径的情况下增强网络学习的能力。

2 MMD-YOLOv7 方法

本文通过设计多通道坐标注意力（MCCA）模块、多尺度卷积（MSC）模块、引入多分支模块（DBB），形成本文的主要方法（称为 MMD 方法），然后在 YOLOv7 基础上融合 MMD 方法形成本文的最终模型 MMD-YOLOv7，模型结构如图 2 所示。其中 MCCA 模块嵌入到 Concat 和 E-ELAN 之间，MSC 模块嵌入到 ELAN 模块内部形成新模块 ELAN_MSC，DBB 模块嵌入到 MP 模块中形成新模块 MP_DBB。

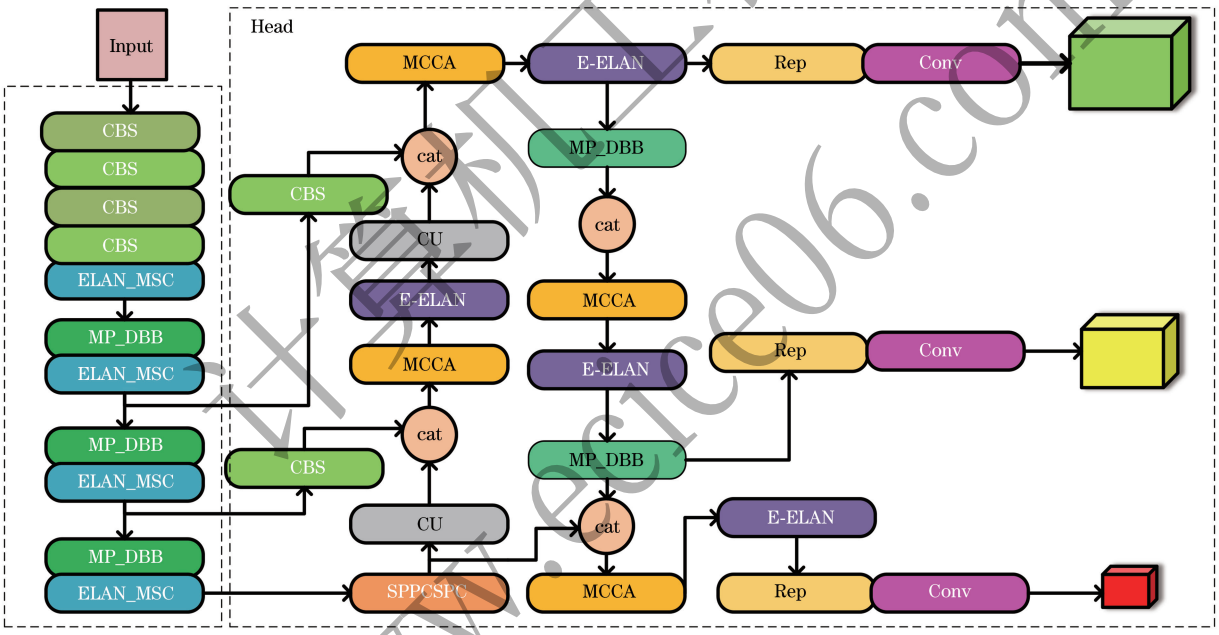


图 2 MMD-YOLOv7 模型的结构

Fig. 2 The structure of MMD-YOLOv7 model

2.1 MCCA 模块设计

CA 是一种高效的注意力机制，其特色是通过关注目标位置信息提高检测性能。但在复杂或低对比度的背景下，CA 不能充分分辨前景目标和背景噪声，极有可能导致重要的上下文信息被忽略。为了解决这个问题，本文以 CA 为蓝本，聚合多通道坐标信息，设计了 MCCA 模块，既充分发挥 CA 的优势，又能帮助模型学习更复杂的空间依赖关系。MCCA 结构如图 3 所示，输入通道数为 C 的特征图 x_c，分别用 H、W 表示其高度维数和宽度维数，$\oplus$ 表示拼接运算，$\odot$ 表示逐元素相乘运算，Split 是将特征图沿不同维度进行分割，Mean 是对特征图沿着合并维度进行平均运算。

首先对输入特征图 x_c 在高度和宽度两个方向上分别进行自适应平均池化，得到高度维上的特征图 z^h 和宽度维上的特征图 z^w，分别如式（1）和式（2）所示：

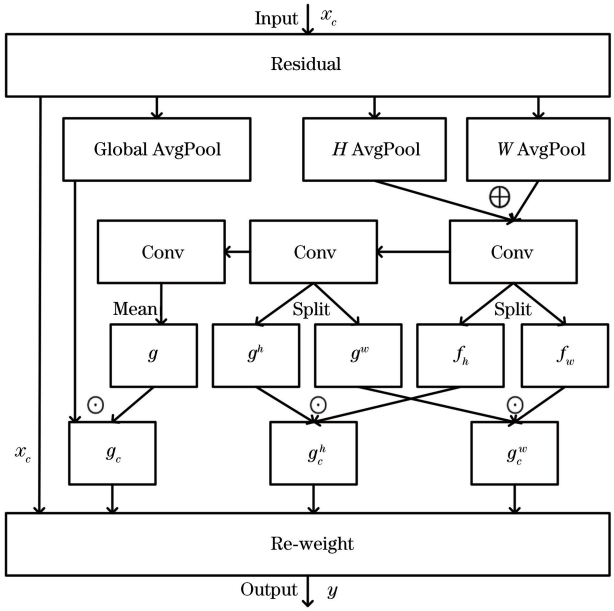


图 3 MCCA 模块结构

Fig. 3 The structure of MCCA module

计算机工程 www.ecice06.com

$$z^h=\frac{1}{W}\sum_{0\leqslant i<W}x_c(h,i) \tag{1}$$

$$z^w=\frac{1}{H}\sum_{0\leqslant j<H}x_c(j,w) \tag{2}$$

将高度和宽度方向的特征图沿通道轴拼接，并通过卷积层处理得到一个综合了高度和宽度全局信息的特征图 f，如式(3)所示：

$$f=\mathrm{Conv}(z^h\oplus z^w) \tag{3}$$

融合的特征图 f 按照原来的高度和宽度进行分割，从而分别得到通道数与原来一样的特征图 f_h 和 f_w。特征图 f 应用另一卷积层处理后，沿着合并维度进行平均池化得到全局平均权重 g，同时进行两个方向的分割从而得到注意力权重 g^h 和 g^w，将分割的权重分别应用到高度和宽度方向的特征上，g_c^h 和 g_c^w 的计算公式分别如式(4)和式(5)所示：

$$g_c^h=f_h\odot g^h \tag{4}$$

$$g_c^w=f_w\odot g^w \tag{5}$$

同时，初始输入特征图 x_c 通过全局平均池化，并对高度和宽度进行压缩，生成新的特征图 f_c，再将全局平均权重 g 应用到新特征图 f_c 上得到 $g_c=f_c\odot g$。最后，将 g_c 与 g_c^h、g_c^w 对初始输入特征图 x_c 联合赋权，从而得到最终的加权特征图 y 作为 MCCA 模块的输出，计算公式如式(6)所示：

$$y=x_c\odot g_c\odot g_c^h\odot g_c^w \tag{6}$$

为了提高模型对复杂场景的识别准确率，将 MCCA 嵌入到 E-ELAN 结构之前(如图 2 所示)，使得特征在进入 E-ELAN 结构之前经过预处理，突出关键位置的特征，强化模型对有限特征的感知能力，提高对灰暗环境的适应性。

2.2 MSC 模块设计

为了应对夜间图像噪声问题，提升模型处理夜间图像的能力，构建多尺度卷积(MSC)模块，用于替换ELAN结构中的部分CBS。MSC模块结构如图 4 所示，其中 C 表示通道数。

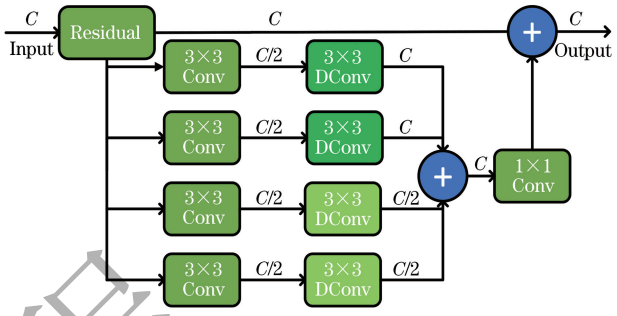


图 4 MSC 模块结构

Fig.4 The structure of MSC module

MSC 模块首先使用 3×3 的卷积层，将输入特征的通道数减半。接下来将处理后的 4 个卷积分别由两个扩张率为 3 和 2 个扩张率为 5 的 3×3 空洞卷积进行处理，并将扩张率为 5 处理过的两组卷积的通道数恢复。空洞卷积是一种扩大卷积核感受野的技术，不同扩张率的卷积层可以捕获不同尺度的空间信息。感受野 F 是特征图上的像素点映射回输入图像上的区域大小，计算公式如式(7)所示：

$$F=k+(k-1)(r-1) \tag{7}$$

式中：k 表示卷积核大小；r 表示扩张率。

最后，将 4 个卷积进行特征融合后，再使用 1×1 卷积与残差连接的结合来增强特征传递并减少信息损失。残差连接允许输入数据直接跳过一些层传递到更深的层，有助于解决深层网络中的梯度消失问题。

通过引入减半通道数策略和空洞卷积，MSC 模块在不增加额外计算负担的情况下，有效地捕获了从近处细节到远处背景的多尺度空间信息。应用到 ELAN 模块得到新的 ELAN_MSC 模块，如图 5 所示，增强模型对夜间图像的处理效率，同时减少信息损失提升特征传递效率，使模型在复杂的夜间场景下仍能保持高准确性。

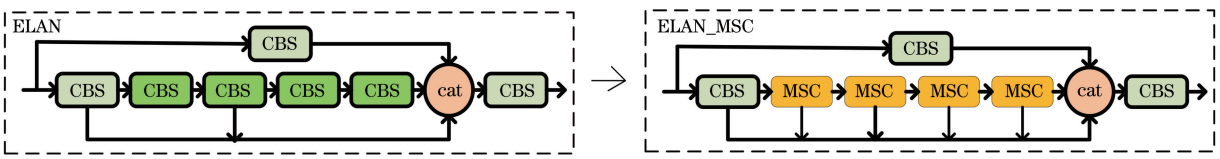


图 5 ELAN 模块和 ELAN_MSC 模块的结构

Fig.5 The structure of ELAN module and ELAN_MSC module

2.3 DBB 模块的引入

YOLOv7 的 MP 模块通过 MaxPool 提取局部区域的最大值信息，通过 CBS 捕获局部区域的所有值信息，提升了网络对特征的感知能力。DBB 汲取了 Inception 模型设计理念之精髓，通过其动态多分支结构显著增强了特征提取能力。而且，DBB 能够等价转换为单个卷积层，可以无缝嵌入任意架构中而不会降低推理速度。本文在 YOLOv7 模型中引入 DBB，通过强强联合进一步增强模型性能。

DBB 采用多分支拓扑结构，通过多种方式增强传统卷积，例如使用 1×1 卷积、1×1 后接平均池化、1×1 后接 $K\times K$ 卷积等，从而可以自由组合成

计算机工程 www.ecice.com

多种不同状态的结构，本文在 6 种常用的 DBB 结构（如图 6 所示）中通过比较实验进行模型选择。首先将这 6 种结构分别嵌入 YOLOv7 的 MP 模块中，嵌入方式如图 7 所示，然后进行 6 种不同 DBB 结构的对比实验，得到实验结果如表 1 所示，其中加粗数字表示最优值。

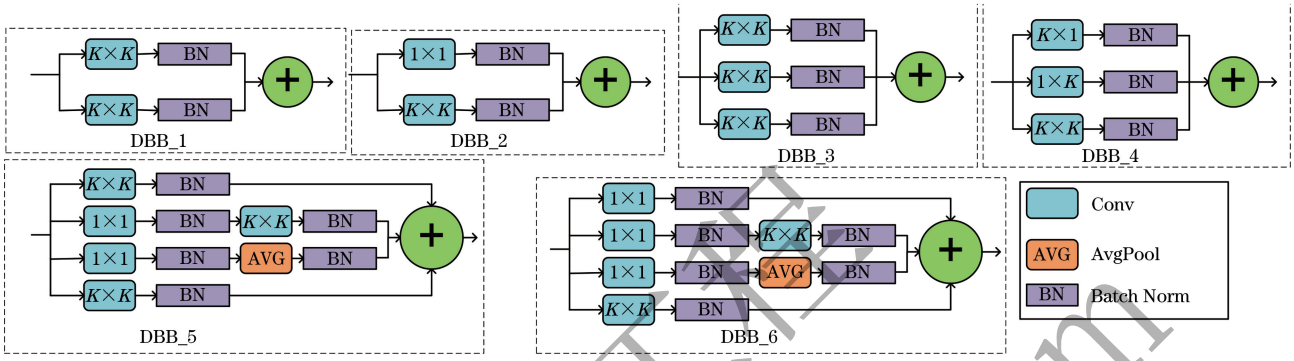


图 6 DBB 模块结构

Fig. 6 The structure of DBB module

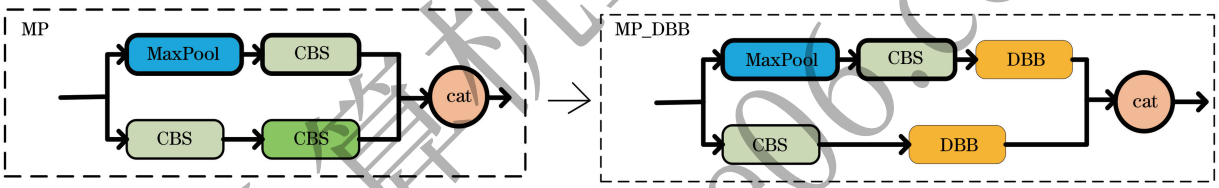


图 7 MP 模块和 MP_DBB 模块的结构

Fig. 7 The structure of MP module and MP_DBB module

表 1 不同结构的 DBB 性能对比

Table 1 Performance comparison of DBB with different structures

Algorithm	Model	mAP@0.5/%
YOLOv7	—	52.6
YOLOv7	DBB_1	51.7
YOLOv7	DBB_2	51.5
YOLOv7	DBB_3	52.3
YOLOv7	DBB_4	52.2
YOLOv7	DBB_5	52.6
YOLOv7	DBB_6	**53.0**

基于实验结果，本文采用 DBB_6 模块嵌入到原有的 MP 结构中形成新模块 MP_DBB。

表 2 实验相关的硬件配置和模型参数

Table 2 Hardware configuration and model parameters related to the experiment

Parameter	Configuration
GPU	A100
CPU	Intel Xeon Platinum 8358
CUDA	11.7
Learning Rate	0.01
Epochs/次	200
Momentum	0.937
Loss	0
Batch Size	64
Optimizer	SGD
Image Size/像素	640×640

3 实验结果与分析

3.1 实验环境与参数

本文所使用的实验环境为 Ubuntu 20.04、Python 3.11.0、PyTorch 2.0.1，相关硬件配置及模型参数如表 2 所示，其他未提到的参数均与 YOLOv7 模型的基本配置相同。

3.2 数据集

BDD100K 数据集由加州大学伯克利分校的人工智能（AI）实验室（BAIR）发布，是目前公开的驾驶数据集中规模最大、内容最多样的数据集之一。该数据集是在现实生活中的多种场景中收集，涵盖了不同的天气、驾驶环境以及一天中不同的时间段。BDD100K 数据集包含了 100 000 张图像，每张图像都进行了详尽的标注。本文中使用的数据集是从 BDD100K 中抽取了 6 000 张黑暗或低光照下的图片，并重新标注，将这些图像的标签分为 5 类，即汽车、公交车、人、卡车和交通标志。数据集按 8∶2的比例划分，其中训练集中有 5 000 张图像，测试集中有 1 000 张图像。

3.3 评价指标

平均精度（AP，R_{AP}）指标是通过计算精度召回（PR）曲线下的面积得到的，用于衡量某一类别分类

计算机工程 www.ecice06.com

器的性能。这个指标反映了分类器在精确度和召回率之间的权衡关系。精确度(Precision,P)和召回率(Recall,R)的计算公式如式(8)和式(9)所示:

$$P=\frac{N_{TP}}{N_{TP}+N_{FP}} \tag{8}$$

$$R=\frac{N_{TP}}{N_{TP}+N_{FN}} \tag{9}$$

式中:N_{TP} 表示模型正确将正样本预测为正样本的数量;N_{FP} 表示模型错误将负样本预测为正样本的数量;N_{FN} 表示模型错误将正样本预测为负样本的数量。通过 PR 曲线,可以根据式(10)得到平均精度值:

$$R_{AP}=\int_0^1 P(R)\mathrm{d}R \tag{10}$$

均值平均精度(mAP,P_{mAP})是通过对所有类别的平均精度值取平均得到的,用以评价目标检测模型的总体性能,计算公式如式(11)所示:

$$P_{mAP}=\frac{1}{N}\sum_{j=1}^{N}R_{AP}^{(j)} \tag{11}$$

在本文中使用 mAP@0.5、参数量(Parameters)、浮点运算次数(FLOPs)和每秒处理的图像帧数(FPS)作为评价指标。这些指标分别代表在交并比(IoU)阈值为 0.5 时的类别平均精度、模型的内存占用量、模型复杂度以及模型处理速度,提供了一个全面的视角来评估和比较不同的目标检测模型。其中,mAP@0.5 尤为重要,因为它直接反映了模型对实际应用场景中对象检测的准确性和鲁棒性。

3.4 对比实验结果

在进行目标检测模型的选择过程中,本文对不同的模型在该数据集上的性能进行了细致比较。这一比较过程包括了在数据集上训练和测试各个模型,并以均值平均精度mAP@0.5作为主要的评估指标。基于这些比较结果,决定采用 YOLOv7 模型作为后续实验的基线模型。实验结果表明,在考虑到 FPS 和 mAP@0.5 的综合性能下,基于 YOLOv7 模型进行改进的 MMD-YOLOv7 模型显示出了较其他模型更加出色的性能。相关的对比实验结果如表 3 所示。

表 3 主流目标检测算法模型的对比实验

Table 3 Comparative experiments of mainstream target detection algorithm Models

Model	Parameters/10^6	FLOPs/10^9	mAP@0.5/%	FPS/(帧·s^{-1})
SSD	24.1	61.2	30.5	61
Faster R-CNN	136.8	369.8	48.4	43
RetinaNet	36.1	98.4	42.4	50
Efficientdet_d0	**3.9**	**2.5**	34.7	36
DETR	41.0	225.0	40.2	20
YOLOv5s	20.8	47.9	49.7	119
YOLOv5m	7.0	15.8	45.8	**140**
YOLOV8m	25.8	78.7	46.0	117
YOLOV8l	43.6	165.7	48.6	97
YOLOV8x	68.2	258.5	56.1	86
YOLOv7	36.5	103.2	52.7	118
MMD-YOLOv7 (ours)	46.8	126.8	**57.9**	100

3.5 消融实验

为了验证本文算法的改进效果,在环境和参数设置保持一致的情况下,采用消融实验对各个改进方案进行验证。以 YOLOv7 为基线模型,分别进行改进点消融实验,其中,“✓”表示选择对应的改进点,每个改进点首先单独依次添加到 YOLOv7 的原始模型中进行改进点消融实验。模型 8 是本文算法模型 MMD-YOLOv7,消融实验结果如表 4 所示。

表 4 MMD 方法各模块对性能的影响

Table 4 The impact of each module of the MMD method on performance

Model	YOLOv7	MCCA	MSC	DBB	Parameters/10^6	FLOPs/10^9	mAP@0.5/%	FPS/(帧·s^{-1})
1	✓				**36.5**	103.2	52.6	**118**
2	✓	✓			38.5	107.4	54.4	114
3	✓		✓		41.6	126.4	54.1	114
4	✓			✓	**36.5**	103.2	53.0	**118**
5	✓	✓	✓		46.8	**126.8**	56.8	100
6	✓	✓		✓	38.5	107.4	55.6	106
7	✓		✓	✓	41.6	126.4	55.8	114
8	✓	✓	✓	✓	46.8	**126.8**	**57.9**	100

针对夜间车辆检测的挑战,本文引入改进的 MCCA 模块到 YOLOv7 模型中,显著提升了网络的检测精度,仅增加了 2×10^6 的参数量。同时,MSC 模块虽增加了 4.9×10^6 的参数量,但精度依旧能够得到提高。此外,DBB 的引入增加了训练时间,但在推理时未影响计算效率或帧率,且轻微提升

计算机工程 www.ecice06.com

了模型精度。结合 MCCA 和 MSC 模块能使 mAP@0.5 提升 4.2 百分点，而 DBB 的加入，在不影响计算参数和推理时间的前提下，进一步提升了模型精度。最终，结合这 3 个高效的模块，改进模型在精度上相比原始模型提高了 5.3 百分点，证明了 MMD 方法在夜间车辆检测领域的出色性能和实用价值。

3.6 模型在其他数据集上的对比

为了验证模型在不同场景和小目标检测方面的效果，将本文模型在 UA-DETRAC[24]、BIT-Vehicle[25]、VirdDrone[26]、Exdark 数据集上进行检测，并进行对比实验。其对比实验结果如表 5 所示。

表 5 不同模型在 4 个公开数据集上的性能比较

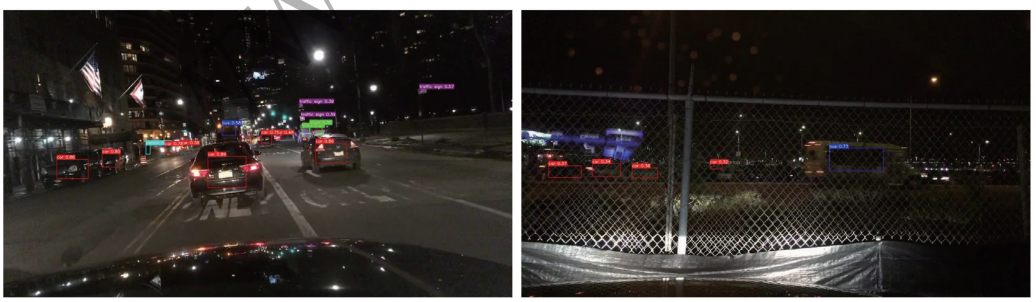
Table 5 Performance comparisons of different models on 4 public datasets

Model	Parameters/10^6	FLOPs/10^9	mAP@0.5/%			
			UA-DETRAC	BIT-Vehicle	VirdDrone	Exdark
SSD	24.1	61.2	39.0	79.2	22.8	41.3
Faster R-CNN	136.8	369.8	47.6	87.2	22.4	35.0
RetinaNet	36.1	98.4	66.3	93.6	25.2	32.2
Efficientdet_d0	**3.9**	**2.5**	43.5	95.4	28.4	34.1
DETR	41.0	225.0	65.0	96.2	30.3	62.2
YOLOv5s	20.8	47.9	57.4	97.3	38.0	58.9
YOLOv5m	7.0	15.8	65.4	96.4	32.0	59.0
YOLOV8m	25.8	78.7	65.8	96.4	41.7	61.8
YOLOV8l	43.6	165.7	68.1	97.3	43.6	63.8
YOLOV8x	68.2	258.5	69.2	97.3	46.3	64.2
YOLOv7	36.5	103.2	63.6	97.3	47.6	68.1
MMD-YOLOv7	46.8	126.8	**71.1**	**97.6**	**49.0**	**70.1**

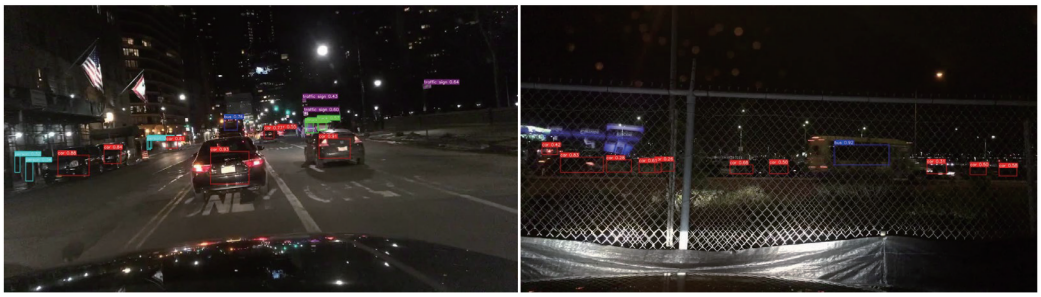
从上述 4 个数据集的对比实验中可以看出，本文提出的改进模型在不同场景、不同天气上的精度都优于原模型，从而进一步验证了本文方法的有效性和可靠性。

3.7 可视化分析

图 8 展示了使用本文提出的 MMD-YOLOv7 模型与 YOLOv7 模型在所采用数据集上的检测结果对比，图 8(a)展示了 YOLOv7 模型的检测结果，图 8(b)则展示了 MMD-YOLOv7 模型的检测结果。图 8 左边的对比图展示了在夜晚光照不足且目标较小的条件下，原模型未能识别到路边的行人，而 MMD-YOLOv7模型成功捕捉到了这些难以察觉



(a)YOLOv7模型



(b)MMD-YOLOv7模型

图 8 在 BBD 数据集上的检测结果对比

Fig.8 Comparison of detection results on BBD dataset

计算机工程 www.ecice06.com

的目标。在图 8 右边对比图中，面临车辆遮挡和重叠的复杂场景，原模型漏检了一些车辆，而 MMD-YOLOv7 模型全面检测到所有车辆。这一对比直观地揭示了 MMD-YOLOv7 模型在提升检测精度方面的显著优势，特别是其在识别微小目标和处理目标重叠情况下的能力。

图 9 展示了模型在不同数据集不同场景下的检测对比结果，从左到右第 1 组图是在城市公路上摄像头所拍的车辆行驶的状况进行的对比，第 2 组图是使用无人机在高速公路上所拍摄高速行驶的车辆进行的对比，第 3 组图是在城市夜间的日常生活进行的对比。从这 3 组对比图中可以明确看出，无论是白天还是黑夜，MMD-YOLOv7 模型对于车辆的检测的准确性比原模型要高，并且能够检测出原模型无法检测的车辆，并能正确识别出被遮掩的物体。



(a)YOLOv7模型

(b)MMD-YOLOv7模型

图 9 在多种场景下的检测结果对比

Fig. 9 Comparison of detection results in various scenarios

综上所述，这些多种场景下的对比结果不仅展示了 MMD-YOLOv7 模型在特定场景下的性能，更验证了其在不同环境下的普适性和优越性。

4 结束语

本文基于 YOLOv7 模型提出一种改进的车辆目标检测的方法，专门针对低光照条件下的车辆检测问题进行优化。首先通过改进的 MCCA 模块来增强模型对全局和局部信息的捕捉能力，从而提升在复杂场景中的识别精度。此外，为了进一步提升模型对小目标的检测效果，本文构建了 MSC 模块。DBB 的引入有效提升了特征提取网络的性能。实验结果表明，与原始 YOLOv7 模型相比，本文提出的改进模型在 BDD 数据集上的精度提升了 5.3 百分点，并在多个公开车辆检测数据集上展示了其优越性能。在探索如何进一步提升模型在极低光照条件下的识别精度上，尽管现有的车辆检测模型已取得显著进展，但在夜间或光照不足的环境中，尤其是当小目标数量众多时，其识别能力仍面临挑战。为了克服这些限制，未来的研究可以从以下方向着手：1）针对低光照条件优化的图像预处理技术，开发专为低光照条件设计的图像预处理技术，如调整图像的亮度、对比度和噪声水平，以改善输入数据的质量，以在模型训练之前提高图像的可用性，从而提高最终的检测精度；2）使用深度学习中的增强学习算法，模型可以在训练过程中逐渐学习如何在各种光照条件下更准确地识别车辆，通过模拟不同光照条件下的场景，使模型能够在实际应用中遇到低光照条件时自我适应并提高识别率；3）利用多种传感器数据与视觉数据的融合，以增强模型在恶劣光照条件下的性能。整合不同类型的感知信息，有助于模型在缺乏可见光信息的情况下依然保持较高的识别准确性。

参考文献

[1] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2023: 7464-7475.

[2] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2021: 13713-13722.

[3] DING X H, ZHANG X Y, HAN J G, et al. Diverse branch block: building a convolution as an inception-like unit[C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2021: 10886-10895.

[4] YU F, CHEN H F, WANG X, et al. BDD100K: a diverse

driving dataset for heterogeneous multitask learning[C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2020: 2636-2645.
[5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2014: 580-587.
[6] GIRSHICK R. Fast R-CNN[C]// Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2015: 1440-1448.
[7] 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, 2017, 39(6): 1137-1149.
[8] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]// Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2017: 2961-2969.
[9] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2020: 10781-10790.
[10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[M]. Berlin, Germany: Springer, 2016.
[11] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2017: 2980-2988.
[12] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 7263-7271.
[13] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. [2023-11-20]. https://arxiv.org/abs/1804.02767v1.
[14] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]// Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2020: 213-229.
[15] HNEWA M, RADHA H. Object detection under rainy conditions for autonomous vehicles: a review of state-of-the-art and emerging techniques[EB/OL]. [2023-11-20]. https://arxiv.org/abs/2006.16471v4.
[16] NABATI R, QI H R. RRPN: radar region proposal network for object detection in autonomous vehicles[C]//Proceedings of IEEE International Conference on Image Processing. Washington D. C., USA: IEEE Press, 2019: 3093-3097.
[17] CARRANZA-GARCÍA M, TORRES-MATEO J, LARA-BENÍTEZ P, et al. On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data[J]. Remote Sensing, 2021, 13(1): 89.
[18] BENJUMEA A, TEETI I, CUZZOLIN F, et al. YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles[EB/OL]. [2023-11-20]. https://arxiv.org/abs/2112.11798v4.
[19] 刘航博, 马礼, 李阳, 等. 无人驾驶中运用 DQN 进行障碍物分类的避障方法[J]. 计算机工程, 2024, 50(11): 380-389.
LIU H B, MA L, LI Y, et al. An obstacle avoidance method using DQN to classify obstacles in unmanned driving[J]. Computer Engineering, 2024, 50(11): 380-389. (in Chinese)
[20] 李松江, 耿兰兰, 王鹏. 基于改进 Yolov4 的车辆目标检测[J]. 计算机工程, 2023, 49(4): 272-280.
LI S J, GENG L L, WANG P. Vehicle target detection based on improved Yolov4[J]. Computer Engineering, 2023, 49(4): 272-280. (in Chinese)
[21] 谢竞, 邓月明, 王润民. 改进 YOLOv8s 的交通标识检测算法[J]. 计算机工程, 2024, 40(11): 338-349.
XIE J, DENG Y M, WANG R M. Improved YOLOv8s traffic sign detection algorithm[J]. Computer Engineering, 2024, 40(11): 338-349. (in Chinese)
[22] 杨秀娟, 曾智勇. 基于 YOLOv5 的无人机航拍改进目标检测算法 Dy-YOLO[J]. 福建师范大学学报(自然科学版), 2024, 40(1): 76-86.
YANG X J, ZENG Z Y. Improved target detection algorithm Dy-YOLO for UAV aerial photography based on YOLOv5[J]. Journal of Fujian Normal University(Natural Science Edition), 2024, 40(1): 76-86. (in Chinese)
[23] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 2117-2125.
[24] WEN L Y, DU D W, CAI Z W, et al. UA-DETRAC: a new benchmark and protocol for multi-object detection and tracking[J]. Computer Vision and Image Understanding, 2020, 193: 102907.
[25] DONG Z, WU Y W, PEI M T, et al. Vehicle type classification using a semisupervised convolutional neural network[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 16(4): 2247-2256.
[26] ZHU P F, WEN L Y, BIAN X, et al. Vision meets drones: a challenge[EB/OL]. [2023-11-20]. https://arxiv.org/abs/1804.07437v2.

编辑 索书志

计算机工程 www.ecice06.com