具有针对性的面向对象优化；②目前所获取的遥感影像空间分辨率虽然较高,但以车辆为代表的小尺寸目标依然无法获取到信息量丰富的视觉特征,致使车辆等小尺寸目标检测精度受限；③遥感影像下小尺寸目标特征不明显、密集车辆目标检测效果较差,难以对具有存在旋转角度的车辆目标进行准确检测。综上所述,现有深度卷积神经网络难以实现“一对一”的车辆目标检测,在遥感影像智慧出行中这一问题尤显突出。

本文主要面向高分辨率遥感影像车辆检测的深度学习模型进行综述并对其适应性开展研究,对目标检测领域的主流算法进行分类,阐述并分析现有深度学习模型应用于遥感影像车辆检测中的优缺点；其次,基于公开数据集,利用主流深度学习模型对遥感影像进行训练,并评估其车辆检测性能；最后,为大幅面、复杂背景环境的小目标车辆检测提供新的解决途径及发展方向。

## 1　基于遥感影像的双阶段目标检测算法

双阶段目标检测算法是在生成候选区域的基础上,对候选区域进行分类与回归,从而得到检测结果。目前,此类算法主要以区域卷积神经网络(region - based convolutional neural network, R - CNN)[7]系列为主,其特点是检测精度较高,但检测速度略显不足。R - CNN 首次将卷积神经网络引入目标检测,结合选择性搜索算法生成候选区域,利用深度卷积神经网络进行特征提取[8-9],在 PASCAL VOC2007 数据集上的目标检测精度为 58.5%。Girshick[10]针对 R - CNN 存在重复计算、检测速度慢的问题提出了 Fast - RCNN 算法,加入感兴趣区域池化模块,使图像输入尺寸不受限制,利用多任务损失函数统一目标分类和候选框回归任务,提高检测的速度。此后,Faster - RCNN 利用区域生成网络代替 Fast - RCNN 中选择性搜索方法,使得候选框数目从原有的约 2 000 个减少为 300 个,提高了检测速度,实现了端到端的目标检测[11]。近几年,遥感领域基于双阶段目标检测算法也开展了系列研究工作,本文将遥感领域双阶段目标检测研究成果的优缺点及检测精度进行总结,如表 1 所示。总结发现,遥感影像目标检测仍存在目标过小、分布密集、角度多样化以及检测背景复杂 4 个难点问题,本文基于遥感影像双阶段车辆检测算法存在的难点问题开展了分析与讨论。

**表 1　双阶段遥感影像车辆目标检测对比**

**Tab. 1　Comparison of two - stage remote sensing image vehicle target detection**

| 模型 | 优点 | 缺点 | 检测精度 |
|---|---|---|---|
| GoogLeNet[12] | 引入超分辨率重建算法加强车辆特征信息 | 对密集目标检测效果差 | 自制遥感影像数据集上精度为 0.75 |
| AVPN + VALN[13] | 提出 VALN 网络实现车辆方向信息检测,融合深浅层特征信息以提高小目标检测精度 | 网络结构复杂,训练检测时间成本较高 | DLR[14]数据集上精度为 0.92 |
| ZF[15] | 针对数据集自身特点,结合区域生成网络设置 3 种对应大小及比例的锚框,加快了检测速度 | 数据量小、来源单一,模型检测鲁棒性较差 | 自制遥感影像数据集上精度为 0.88 |
| DF - RCNN[16] | 融合深浅层特征信息、引入可变形卷积和可变形感兴趣区域池化,改善密集区域小目标检测效果 | 网络计算量较大,检测耗时较长 | 自制 Google Earth 数据集上精度为 0.94 |
| VGGNet[17] | 网络层数可以随数据集图像大小进行调整 | 图像深层特征提取不充分,没有充分利用上下文语义信息 | 自制 Google Earth 数据集上精度为 0.89 |
| VGG - 16[18] | 基于超像素分割提取道路区域进行车辆检测,缩小了检测范围 | 操作较为复杂,检测时间较长,对于非道路区域内车辆会造成漏检 | DOTA[19]数据集上精度为 0.73 |
| SORCN[20] | 加入道路区域分割提高检测精度 | 运行速度较慢,无法预测车辆方向信息,模型抗干扰性差,对于非道路区域内车辆无法检测 | 自制 Google Earth 数据集上精度为 0.96 |
| Faster R - CNN + +[21] | 对遥感影像车辆进行多尺度融合及数据增强,提高了模型鲁棒性与小目标检测精度 | 检测速度较慢,无法实现车辆方向检测 | DLR 数据集上精度为 0.57；Potsdam[22]数据集上精度为 0.67；VEDAI[23]数据集上精度为 0.458 |

1)基于遥感影像的小目标检测。鉴于遥感影像中车辆目标尺寸较小,多次卷积将导致特征信息丢失严重,造成较大程度的漏检问题,其主要改进思路有：区域预提取、优化锚框策略和改进网络结构等。对遥感影像进行道路区域预提取,再将道路区域输入卷积神经网络识别车辆目标,可提高车辆目

标的检测率[18]，该算法区域预提取阶段使用的是传统方法，道路区域提取的自动化能力不足，对非道路车辆信息无法进行检测。SORCN（small object recognition convolutional network）[20]是一种基于分割[24]的小物体识别网络，该模型通过分割图像车辆目标，进一步提高小目标车辆检测效果。该方法可以减小待检测区域、提高车辆小目标检测精度，但流程较为复杂，需要牺牲检测速度提高检测精度。基于锚框的目标检测算法，锚框大小及比例选取直接影响待检测目标的召回率与检测精度。利用聚类分析方法计算出适合训练集的锚框尺寸，降低目标定位难度，提高小目标检测精度[15]。但该方法泛化能力较弱，难以应用于差异较大的数据集。关于网络结构的优化[17,25]，顾及小目标特征信息丢失问题，主要通过减少网络下采样次数增加特征图尺寸，但会引入深层高级语义特征信息提取不足的问题。

2）基于遥感影像的多角度小目标检测。遥感影像具有多角度、多传感器、多分辨率的特点，车辆目标方向往往具有不确定性，直接使用水平方向检测框，使得较多非目标干扰信息介入而影响检测精度。针对多角度小目标检测问题，可通过旋转扩充增强[26-27]、引入旋转区域生成网络来改善多角度小目标检测效果。旋转扩充增强检测到的角度信息有限，利用特定旋转区域建议网络生成预选框，可实现对任意角度信息的预测。$R^3$ - Net[28]和 $R^2$PN[29]均是基于区域建议网络融合生成旋转候选区域，实现车辆角度信息预测，但网络计算量也相应增加。对于遥感影像目标角度问题，目前主要通过引入旋转区域生成网络来进行角度预测，该方法可缓解角度问题对检测精度的影响，但网络会变得更加复杂，增加了目标检测时间。

3）基于遥感影像的密集型小目标检测。遥感影像中密集型车辆检测难点主要在于车辆目标尺寸太小并且具有旋转角度，导致密集型目标存在漏检现象，现有方法多通过提高小目标检测精度、增加角度预测模块来减小密集目标所带来的检测影响。DF - RCNN（Deformable Faster - RCNN）模型[16]融合深浅层特征信息提高密集型小目标的检测效果。PVANet 模型引入角度检测模块，减少预测结果间重合度，以减少对密集排列目标的漏检，该方法对密集目标的定位效果较好，但难以解决单个车辆目标提取的问题[30]。

4）基于复杂背景的遥感影像小目标检测。遥感影像自身具有覆盖范围广、地物类别复杂多样、成像受到云层等因素干扰的特点，给车辆精准识别带来了挑战。鉴于遥感影像复杂背景下小目标检测的难点问题，现有研究成果主要从引入区域预提取与注意力机制模块这 2 个方面来解决。对图像进行去雾、除云等预处理操作，可减少干扰信息影响，但此类预处理操作会降低影像分辨率。通过超分辨率重建算法可增强遥感影像特征[12]，在不损失影像分辨率前提下提高检测精度。但该方法会加大网络复杂程度及计算量，难以满足实时性需求。基于注意力机制的特征融合可以减弱背景信息的干扰，改善复杂背景下小目标的检测效果[31]。由此可见，采用区域预提取以及注意力机制可以减小背景信息的干扰，但区域预提取时间成本较大，注意力机制将是未来解决复杂背景及噪声干扰问题的重要研究方向。

综上所述，目标固有尺寸过小是限制遥感影像车辆检测算法性能的首要因素，集中解决目标检测中存在的小尺度、多角度、密集型以及复杂背景干扰等因素是当前遥感影像车辆检测的首要任务，需综合考虑各个问题之间存在的内在关联性，借鉴解决不同问题的优化策略，以期达到最终提高检测性能的目标。

## 2 基于遥感影像的单阶段目标检测算法

单阶段目标检测无需生成候选区域，隶属于基于回归分析思想的检测算法。主流单阶段目标检测算法有 YOLO（you only look once），SSD（single shot multibox detector）[32]以及无锚框目标检测系列，此类算法的特点是检测速度较快，但检测精度相对于双阶段目标检测略有不足。

YOLO 系列作为单阶段目标检测的主流算法，2016 年 Redmon 等[33]首次提出 YOLO 模型，将检测转化为回归问题，直接对目标进行定位和类别预测。在此基础上，YOLOv2 引入锚框机制及特征尺度融合模块，改善了定位精度以及小目标检测能力的不足[34]。YOLOv3 利用 Darknet - 53 骨干网络增强特征提取能力，设计 3 种不同尺度网络预测目标，多尺度目标检测能力更强[35]。YOLOv4 使用 CSPDarknet 作为特征提取网络，融合 SPP - Net 和 PANet 提高检测精度和速度，使检测器在单个 GPU 上也能很好地完成训练[36]。YOLOv5 与 YOLOv4 结构相似，网络模型更加轻量，训练速度远超 YOLOv4，兼顾速度的同时保证了准确性。

SSD 模型应用多尺度特征进行目标检测，借鉴 Faster R - CNN 中锚框的理念，设置不同尺度与长宽比的锚框，相对于 YOLOv1 检测效果有明显提升[37]。在 SSD 的基础上，DSSD[38]算法改用 Res-

Net101 作为特征提取网络,利用反卷积传递深层特征,融合深浅层特征信息,从而提高了对小目标的检测效果。

无锚框目标检测算法通过检测中心点或关键角点进行目标边界框的预测,无需设定锚框。FCOS (fully convolutional one - stage object detection)[39]是一种无锚框的单阶段全卷积目标检测算法,对特征图进行像素级回归,直接预测目标中心点、边框距中心点的距离来检测目标。无锚框的目标检测算法具有更灵活的解空间,不需要调优与锚框相关的超参数,极大减少了算法计算量,训练过程内存占用更低。

相比于双阶段目标检测算法,单阶段目标检测的精度较低,但速度较快。本文将遥感领域现有单阶段目标检测研究成果进行归纳总结,如表 2 所示。基于遥感影像的单阶段车辆检测算法研究较少,总结过程中也参考了一些相关研究文献,包括飞机、船舰等,并对遥感影像下目标小、密集排列等问题的优化及解决思路进行了归纳总结。

**表 2　单阶段遥感影像车辆目标检测对比**

**Tab. 2　Comparison of one - stage remote sensing image vehicle target detection**

| 模型 | 优点 | 缺点 | 检测精度 |
|---|---|---|---|
| YOLO[40] | 基于 K - means 算法重新计算锚框大小,提高小目标检测精度 | 数据来源单一,难以应用于其他类型数据 | 自制 Google Earth 数据集上精度为 0.95 |
| YOLT[41] | 对遥感影像进行有重叠区域的裁剪,确保目标信息完整性 | 裁剪后的图像具有重叠区域,存在较多冗余计算 | 自制遥感影像数据集上精度为 0.9 |
| YOLOv3[42] | 减少特征提取网络层数,增加精细输出特征图来检测小目标 | 数据集规模较小,鲁棒性较差 | 自制高分二号卫星影像数据集上精度为 0.743 |
| Oriented_SSD[43] | 引入目标的角度偏移量,实现车辆角度信息预测 | 虽实现了角度信息的预测,但精度较低 | DLR 数据集上精度为 0.86; VEDAI 数据集上精度为 0.8 |

1)基于遥感影像的小目标检测。单阶段目标检测算法主要从优化锚框策略、融合不同层级特征信息和增加网络检测头 3 个方面提高检测精度。其中,优化锚框策略可以提高对多尺度目标的检测能力,是当前常见的优化方法[40,44],但过于依赖先验设计,难以应用于其他类型数据。融合不同层级特征,可以兼顾浅层特征的纹理、边缘等细节信息,优化网络结构的同时使得输出特征图的尺度更适合小目标检测,改善小目标检测能力[41,45],但融合浅层特征的目标检测精度提升有限,在小而密集的目标场景中会存在较多漏检现象。在深度学习目标检测模型中增加网络检测头,如在 YOLOv3 网络增加大小为 104 像素 × 104 像素检测头[42],用于检测小尺度目标,可以提高小目标检测性能,但该方法会增加网络复杂程度、训练及检测时间。

2)基于遥感影像的多角度小目标检测。单阶段目标检测主要通过引入角度因子与优化损失函数 2 个方面来解决目标存在的角度问题。在算法预设锚框中加入角度因子[43,46],可对车辆位置与角度信息同时预测,但加入角度参数会增加网络计算复杂度[47]。R - YOLO 模型引入新的损失函数和旋转交并比计算方式,实现了舰船目标任意角度预测[48]。上述方法均引入了额外的计算模块,算法复杂度有所增加,在牺牲算法效率的基础上,提高了对多角度小目标的检测能力。

3)基于遥感影像的密集型小目标检测。现有的解决思路包括简化网络结构、增加角度预测分支与改进损失函数。简化特征网络提取更多浅层特征,利用残差网络代替连续卷积减少梯度消失,可提高小目标检测能力,从而减少密集型车辆的漏检数量[49]。该方法未消除目标方向多样所带来的干扰,存在较多重叠检测而导致的目标漏检。在网络中加入角度回归分支,引入非对称卷积增强目标旋转不变性特征,可缓解目标角度对检测精度的影响[50]。损失函数对密集目标定位效果有着重要影响,基于 YOLOv5 算法,利用 CIOU_LOSS 损失函数代替 GIOU_LOSS 损失函数,可减少密集区域的小目标漏检[51]。

4)基于复杂背景的遥感影像小目标检测。针对遥感影像背景复杂及噪声干扰问题,主要解决思路有影像预处理、融合多源影像信息以及添加注意力机制 3 种。影像预处理,如引入基于暗通道先验的大气校正方法,减少大气吸收和散射对遥感影像的影响,可减弱噪声干扰对检测造成的影响[52],但此类方法会影响影像自身信息量。融合可见光、合成孔径雷达遥感影像等多源图像信息,充分利用多源图像各自的成像优势,在复杂背景下能获得更好的检测效果[48]。以上方法可一定程度提高模型的抗干扰性,但相对比较耗时,且对多种复杂的噪声干扰敏感性较弱。在特征融合阶段引入注意力机制能使网络提取更多重要特征信息,充分挖掘小目标的上下文语义特征信息,抑制无关信息的干扰[45]。注意力机制与网络结合方式多样,合理选择融入方式,能有效降低背景信息对检测精度的影响。

由此可见,现有遥感影像车辆检测算法已取得了一定的进步,但仍有较大的提升空间,距离实际工程化应用还尚有距离,有待后续持续的深入研究。

## 3　实验结果及分析

为分析主流深度学习模型在遥感数据集的车辆检测效果,实验分别选取双阶段的 Faster - RCNN,单阶段的 SSD,FCOS 和 YOLOv5 进行测试,其中 FCOS 为无锚框目标检测算法。另外,针对小尺度、密集型、多角度、复杂背景和动态区域 5 种不同场景进行车辆检测测试,分析不同算法在不同场景下的目标检测性能。

### 3.1　数据集及实验设置

实验采用 DOTA 和 DIOR[53] 这 2 个数据集进行测试。其中,DOTA 数据集共有 2 806 张遥感影像和 188 282 个标注实例,图像大小在 800 像素 ×800 像素到 4 000 像素 ×4 000 像素之间,空间分辨率为 0.1 ~4.5 m,包含车辆、飞机、储油罐和游泳池等 15 种标注类别,本文展示的 5 张影像数据空间分辨率为 0.1 ~0.2 m。DIOR 数据集共有 23 463 张遥感影像和 190 288 个标注实例,图像大小为 800 像素 ×800 像素,空间分辨率为 0.5 ~30 m,包含车辆、飞机、机场和棒球场等 20 种目标,鉴于原 DIOR 数据集未注明每张影像空间分辨率信息,因此未列出本文展示影像数据的空间分辨率信息。

本实验硬件设备配置:操作系统为 Ubuntu16.04,GPU 型号为 GeForce1080 Ti(11 G),CPU 型号为 Intel core i9 - 9900K。FCOS,Faster - RCNN,SSD 和 YOLOv5 特征提取网络分别选用 ResNet50 - FPN,ResNet50,MobileNet 和 CSP - Darknet53。实验模型进行 100 次迭代训练,初始学习率设置为 0.01,批处理大小设置为 8,在进行至 80 次训练后学习率调整为原学习率的 10%,进行至 90 次训练后学习率调整为原学习率的 1%,其他参数与模型的官方代码参数设置保持一致。

### 3.2　评价指标

本文通过定性和定量 2 种评价方式对各算法进行综合评定,定量评价使用交并比(intersection over union,IOU)、准确率、召回率、平均准确率(average precision, AP)作为单类别目标检测效果的评价指标,其中平均准确率是评价模型精度的最常用的指标,它是准确率 - 召回率(precision - recall, P - R)曲线下所围成的面积,通常 AP 值越高,模型检测效果越好。4 个指标的计算公式分别为:

$$IOU = \frac{|A \cap B|}{|A \cup B|}, \tag{1}$$

$$P = \frac{TP}{TP + FP}, \tag{2}$$

$$R = \frac{TP}{TP + FN}, \tag{3}$$

$$AP = \int_0^1 P(R)\,\mathrm{d}R, \tag{4}$$

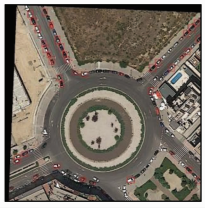
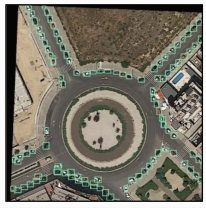
式中:$A$ 为真实框面积;$B$ 为预测框面积;$TP$ 为被正确识别的正样本数量;$FP$ 为被错误识别的正样本数量;$FN$ 为被错误识别的负样本数量;$IOU$ 为交并比;$P$ 为准确率;$R$ 为召回率;$AP$ 为平均准确率。

### 3.3　实验结果分析

1)小尺度车辆目标检测。从表 3 可以看出,在 DOTA 和 DIOR 数据集上 YOLOv5 对于车辆检测均展现了相对较高的性能,漏检目标较少。YOLOv5 算法加入自适应锚框模块,根据不同数据类型计算

**表 3　小尺度车辆目标检测结果对比**

**Tab.3　Comparison of small scale vehicle target detection results**

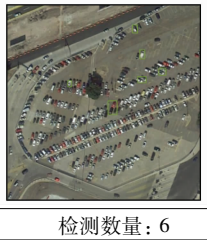
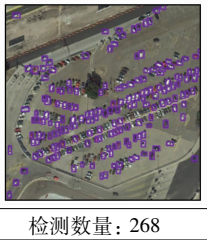
| 影像 数据集 | 目标数量 | SSD 结果 | FCOS 结果 | Faster - RCNN 结果 | YOLOv5 结果 |
|---|---|---|---|---|---|
| 影像 1 DOTA | 111 | |  | |  |
| | | 检测数量:11 | 检测数量:34 | 检测数量:89 | 检测数量:110 |
| 影像 2 DIOR | 86 | | | | |
| | | 检测数量:0 | 检测数量:0 | 检测数量:8 | 检测数量:27 |

最佳锚框大小,使锚框更适合待检测目标。FCOS 和 SSD 算法目标检测精度相对较差,其中 SSD 漏检目标较多。Faster - RCNN 和 SSD 算法使用预设锚框,对于待检测目标尺度大小的自适应性较差,同时 FCOS 和 SSD 算法的特征图感受野相对较大,致使车辆细节特征提取不足,造成大量漏检。结合 2 个数据集检测结果进行对比,DIOR 数据集的检测精度整体低于 DOTA 数据集,其原因是 DIOR 数据集车辆目标小于 5 个像素,更加考验深度学习模型对小目标的检测性能。

2)密集型小尺度车辆目标检测。实验针对静态场景下小而密集型车辆目标进行深度学习模型性能测试,如表 4 所示,在 2 个公开数据集上,YOLOv5 算法的目标检测精度相对较好,但对于具有旋转角度的车辆目标,检测结果间存在较多重叠区域,无法准确定位目标位置。对于密集目标检测结果而言,检测锚框大小与比例的设置尤为重要。Faster - RCNN 锚框相对较大,导致一个检测框内存在多个目标,经过非极大值抑制处理后车辆目标被滤除而产生漏检现象。FCOS 算法对图像进行像素级预测,通过中心点回归目标真实大小,但其浅层特征提取不充分,导致大量车辆漏检。SSD 算法感受野相对较大,检测锚框自适应性较差,相对于其他算法,漏检率相对较高。

**表 4　密集型小尺度车辆目标检测结果对比**

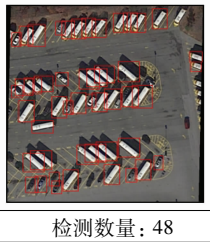
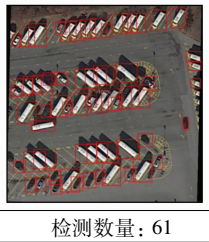
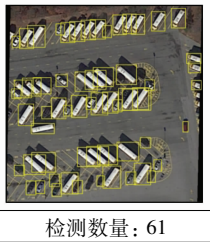
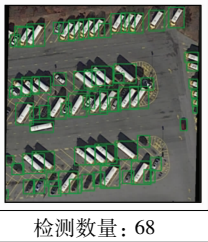
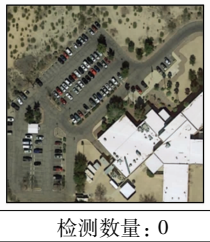
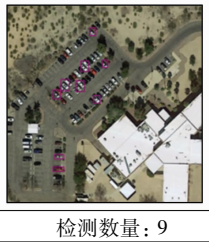
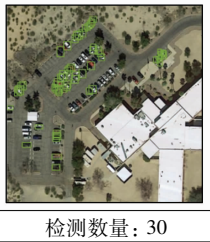
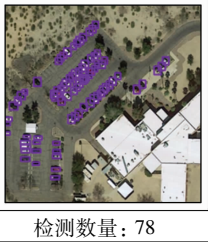
**Tab. 4　Comparison of intensive small scale vehicle target detection results**

| 影像 数据集 | 目标数量 | SSD 结果 | FCOS 结果 | Faster - RCNN 结果 | YOLOv5 结果 |
|---|---|---|---|---|---|
| 影像 3 DOTA | 493 | | | | |
| | | 检测数量: 3 | 检测数量: 95 | 检测数量: 119 | 检测数量: 358 |
| 影像 4 DIOR | 327 | |  | |  |
| | | 检测数量: 0 | 检测数量: 6 | 检测数量: 33 | 检测数量: 268 |

3)多角度小尺度车辆目标检测。由于遥感影像自身特点,车辆目标往往会呈现出多角度排列的状态,从表 5 中可以看出,4 种深度学习模型均可以检测到多角度排列的车辆目标。由于使用水平检测框,无法准确显示角度信息。在车辆密集排列时,目标预测框间会出现不同程度的重叠情况,难以准确

**表 5　多角度小尺度车辆目标检测结果对比**

**Tab. 5　Comparison of multi - angle small scale vehicle target detection results**

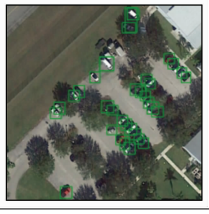
| 影像 数据集 | 目标数量 | SSD 结果 | FCOS 结果 | Faster - RCNN 结果 | YOLOv5 结果 |
|---|---|---|---|---|---|
| 影像 5 DOTA | 68 |  |  |  |  |
| | | 检测数量: 48 | 检测数量: 61 | 检测数量: 61 | 检测数量: 68 |
| 影像 6 DIOR | 78 |  |  |  |  |
| | | 检测数量: 0 | 检测数量: 9 | 检测数量: 30 | 检测数量: 78 |

地进行单目标提取。特别是在 DIOR 数据集中，车辆尺寸进一步缩小时，算法性能下降更加严重。可见在目标密集排列且存在角度时，相邻目标的检测框易出现重叠现象，待检测目标越小，其重叠程度越高，影响了车辆检测精度。

4）复杂背景下小尺度车辆目标检测。实验针对静态场景下复杂背景的遥感影像进行车辆检测，结果如表 6 所示，YOLOv5 的目标检测效果优于其他算法。YOLOv5 加入 Mosaic 数据增强，通过图像扰动、添加噪声和随机缩放裁剪等方式，使算法模型在复杂背景下可以表现出相对较好的检测性能，但依然无法彻底消除噪声干扰。另外 3 种算法仅对数据进行旋转和缩放处理，模型抗噪能力不足，在出现阴影遮挡时会出现较多目标漏检。

**表 6　复杂背景下小尺度车辆目标检测结果对比**

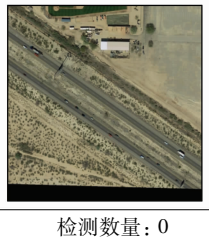
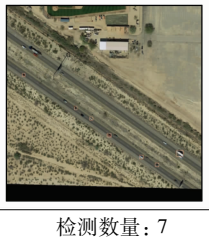
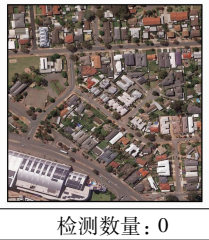
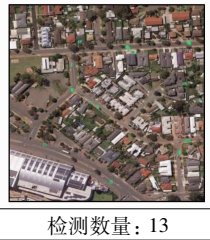
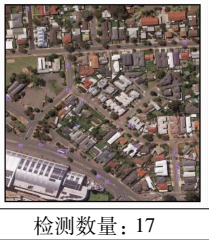
**Tab. 6　Comparison of small scale vehicle target detection results under complex background**

| 影像 数据集 | 目标数量 | SSD 结果 | FCOS 结果 | Faster – RCNN 结果 | YOLOv5 结果 |
|---|---|---|---|---|---|
| 影像 7 DOTA | 40 | | | |  |
| | | 检测数量：6 | 检测数量：4 | 检测数量：30 | 检测数量：32 |
| 影像 8 DIOR | 83 | | | | |
| | | 检测数量：7 | 检测数量：27 | 检测数量：33 | 检测数量：60 |

5）移动小尺度车辆目标检测。相对于以上 4 组静态场景下的车辆检测结果，实验开展了移动场景下车辆检测对比实验，结果如表 7 所示，依然是 YOLOv5 算法的检测效果最好，其次为 Faster – RCNN，FCOS 和 SSD 算法由于小目标检测性能较差，车辆目标漏检较多。主要源于 YOLOv5 算法融合深浅层特征，可获得丰富的全局与局部细节特征信息，对移动场景下的目标检测具有较好的自适应性。在移动车辆场景中，车辆目标分布较为稀疏，拍摄视野更加广阔，车辆小目标特性更加突出，极大考验小目标检测性能。

**表 7　移动小尺度车辆目标检测结果对比**

**Tab. 7　Comparison of moving small scale vehicle target detection results**

| 影像 数据集 | 目标数量 | SSD 结果 | FCOS 结果 | Faster – RCNN 结果 | YOLOv5 结果 |
|---|---|---|---|---|---|
| 影像 9 DOTA | 26 |  |  | | |
| | | 检测数量：0 | 检测数量：7 | 检测数量：16 | 检测数量：24 |
| 影像 10 DIOR | 33 |  | |  |  |
| | | 检测数量：0 | 检测数量：0 | 检测数量：13 | 检测数量：17 |

本文将不同算法在公开数据集上的检测精度及训练时间进行了统计，如表 8 所示，YOLOv5 算法的检测精度最高，在 DOTA 和 DIOR 数据集的检测精度分别为 0.695 和 0.566，SSD 算法检测精度相对较

低,分别为 0.251 和 0.154。主要原因是车辆目标尺度过小且相对密集,该算法对于小尺度目标检测能力相对较差。4 种算法中 SSD 算法训练时间最短,在 DOTA 和 DIOR 数据集上分别为 31.245 h 和 21.135 h,Faster - RCNN 算法训练时间最长,分别为 96.617 h 和 129.617 h,可见单阶段目标检测训练速度优于双阶段目标检测算法。

**表 8　不同算法在公开数据集的检测结果对比**

**Tab. 8　Comparison of detection results of different algorithms in public data sets**

| 模型 | 输入大小/像素 | DOTA | | DIOR | |
|---|---|---|---|---|---|
| | | *AP*① | 训练时间/h | *AP* | 训练时间/h |
| SSD | 800 × 800 | 0.251 | 31.245 | 0.154 | 21.135 |
| FCOS | 800 × 800 | 0.370 | 54.510 | 0.231 | 63.367 |
| Faster - RCNN | 800 × 800 | 0.410 | 96.617 | 0.243 | 129.617 |
| YOLOv5 | 800 × 800 | 0.695 | 43.333 | 0.566 | 38.915 |

①表中 AP 值均为 IOU 阈值大于 0.5 时的 AP 值。

综上所述,YOLOv5 在 DOTA 和 DIOR 数据集的检测效果优于 Faster - RCNN,FCOS 和 SSD 算法。主要原因在于:①YOLOv5 中加入了自适应锚框计算模块,可自动计算最佳锚框类型,提高算法的检测精度和召回率;②YOLOv5 融合不同层级特征,特征信息利用更加充分;③YOLOv5 引入了 Mosaic 数据增强,对图像进行随机缩放、随机裁剪和随机排布等处理,有效提高算法的抗干扰能力。YOLOv5 算法检测速度快,适用于实时性要求较高的场景,但应用于遥感影像车辆检测领域还需要进一步优化,如在遥感影像车辆小而密集的场景下,可能会存在大量小目标车辆漏检,需根据遥感影像车辆目标特点优化网络结构;YOLOv5 算法不具备角度信息检测能力,在车辆密集分布且存在一定角度时,会较大程度地影响其检测精度。

### 3.4　不同数据、模型适应性分析

1)对不同数据的适应性。从实验结果可知,目标检测算法对于 DOTA 数据集的检测精度高于 DIOR 数据集。进一步说明,目标检测算法性能不仅在于模型自身网络结构,也会受数据集大小和内容的影响。DOTA 数据集中的车辆目标数量以及类型相对较多,数据样本较为丰富,所以训练出的模型的鲁棒性更强。因此,高质量、大规模的数据集可以有效提升遥感影像车辆检测效果。

2)对不同模型的适应性。Faster - RCNN 与单阶段目标检测算法不同之处是检测目标时生成候选区域,再对候选区域进行多次分类与位置修正,相对于 FCOS 和 SSD 算法检测速度较慢,但精度较高。然而,Faster - RCNN 锚框类型不适用于小尺度目标,当遥感影像中的车辆目标尺寸更小时,检测效果并不理想。SSD 算法对不同层级选用不同尺寸与比例的锚框,整体精度相对较高,但对小尺度目标检测能力较差,主要原因是特征层非线性化程度不够,模型精度受限。YOLOv5 加入自适应锚框计算与特征融合模块,特征提取能力相对更强,检测性能相对更好,但对于极小尺度的车辆目标,也会出现漏检现象。FCOS 算法无须进行与锚框相关的复杂运算,是一种轻量的检测模型,直接进行逐像素点的回归预测,避免了正负样本不平衡的问题,但其特征图感受野相对较大,难以检测到小尺度车辆目标。

综上所述,对于高精度的车辆检测任务,宜优先使用双阶段的目标检测算法,对于实时性要求高的车辆检测任务,应优先选择单阶段的目标检测算法。现阶段的目标检测算法直接应用于遥感影像车辆检测还难以取得较为理想的效果,网络结构仍有待针对性的改进。

## 4　总结及展望

本文针对主流目标检测模型进行综述,通过 DOTA 和 DIOR 这 2 种数据集进行实验对比,分析不同场景下遥感影像车辆检测效果。对于静态场景而言,车辆排列密集、角度多样,检测结果容易出现预测框重叠现象。而对于动态场景而言,车辆分布相对稀疏,目标角度对于检测结果的影响较小。目前深度学习技术在遥感影像车辆检测任务中已取得了一些成果,但仍存在许多亟待解决的问题,目标过小、分布密集、角度多样以及背景复杂问题是制约遥感影像车辆检测性能的主要因素。因此,针对现存难点问题进行不断优化,是未来遥感影像车辆检测的研究重点,其发展趋势主要涉及以下几点:

1)加强弱监督学习与无监督学习研究。面向遥感影像大数据驱动下的深度学习算法应用中,数据量大,标注耗时长、成本高是数据处理现存的显在问题。如何在弱监督或无监督学习下通过少量标注样本及大量未标注数据进行学习,改善遥感影像标注数据量不足的缺陷,是后续研究亟待解决的问题之一。

2)结合自适应特征尺度融合与注意力机制。鉴于图像高低层特征信息的融合仅将不同分辨率的特征对齐相加,忽略了各特征层之间的关联信息。引入自适应特征融合机制,不同层级特征分配自适应权重参数,实现局部特征和全局特征的高效融合,以期提升车辆小目标检测性能;或者引入注意力机制模块,筛选图像的关键信息,以减少干扰信息对目

标检测的影响,进一步提升小目标检测精度。

3)探索无锚框的目标检测算法。现有研究成果大多是基于锚框进行目标检测,此类算法训练耗时长、运算量大。而无锚框的目标检测算法无须进行大量锚框参数相关的运算,其检测速度更具优势,深入研究将在遥感卫星军事监测及智能交通等领域发挥重要作用。

## 参考文献(References):


[1] 刘金明. 基于深度卷积神经网络的遥感图像中车辆检测方法研究[D]. 开封:河南大学,2020.
Liu J M. Research on vehicle detection method in remote sensing images based on deep convolution neural network[D]. Kaifeng: Henan University,2020.

[2] 刘天颖,李文根,关佶红. 基于深度学习的光学遥感图像目标检测方法综述[J]. 无线电通信技术,2020,46(6):624 - 634.
Liu T Y,Li W G,Guan J H. Deep learning based object detection in optical remote sensing images:A survey[J]. Radio Communications Technology,2020,46(6):624 - 634.

[3] Cheng G, Han J. A survey on object detection in optical remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2016,117:11 - 28.

[4] 成 喆,吕京国,白颖奇,等. 结合 RPN 网络与 SSD 算法的遥感影像目标检测算法[J]. 测绘科学,2021,46(4):75 - 82,99.
Cheng Z,Lyu J G,Bai Y Q,et al. High - resolution remote sensing image object detection algorithm combining RPN network and SSD algorithm[J]. Science of Surveying and Mapping,2021,46(4):75 - 82,99.

[5] Alam M,Wang J F,Cong G,et al. Convolutional neural network for the semantic segmentation of remote sensing images[J]. Mobile Networks and Applications,2021,26:200 - 215.

[6] Ji H,Gao Z,Mei T,et al. Vehicle detection in remote sensing images leveraging on simultaneous super - resolution[J]. IEEE Geoscience and Remote Sensing Letters,2020,17(4):676 - 680.

[7] 张 昭,姚国愉,李雪纯,等. 基于改进 Faster R - CNN 算法的小目标车辆检测[J]. 科技创新与应用,2021(4):28 - 32.
Zhang Z,Yao G Y,Li X C,et al. Small target vehicle detection based on improved Faster - RCNN algorithm[J]. Science and Technology Innovation and Application,2021(4):28 - 32.

[8] Girshick R,Donahue J,Darrell T,et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE,2014:580 - 587.

[9] 南晓虎,丁 雷. 深度学习的典型目标检测算法综述[J]. 计算机应用研究,2020(s2):15 - 21.
Nan X H,Ding L. Overview of typical object detection algorithms based on deep learning[J]. Application Research of Computers, 2020(s2):15 - 21.

[10] Girshick R. Fast R - CNN[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE,2015: 1440 - 1448.

[11] 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.

[12] 罗 峰. 基于超分辨率迁移学习的遥感图像车辆检测[D]. 厦门:厦门大学,2017.
Luo F. Vehicle detection in remote sensing images based on super - resolution transfer learning[D]. Xiamen:Xiamen University,2017.

[13] Deng Z P,Hao S,Zhou S L,et al. Toward fast and accurate vehicle detection in aerial images using coupled region - based convolutional neural networks[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2017,10(8):3652 - 3664.

[14] Liu K,Mattyus G. Fast multiclass vehicle detection on aerial images[J]. IEEE Geoscience and Remote Sensing Letters,2015,12 (9):1938 - 1942.

[15] 王 雪,隋立春,李顶萌,等. 区域卷积神经网络用于遥感影像车辆检测[J]. 公路交通科技,2018,35(3):103 - 108.
Wang X,Sui L C,Li D M,et al. Regional convolutional neural network for vehicle detection in remote sensing images[J]. Journal of Electronics and Information Technology,2018,35(3):103 - 108.

[16] 高 鑫,李 慧,张 义,等. 基于可变形卷积神经网络的遥感影像密集区域车辆检测方法[J]. 电子与信息学报,2018,40 (12):2812 - 2819.
Gao X,Li H,Zhang Y,et al. Vehicle detection in remote sensing images of dense areas based on deformable convolution neural network[J]. Journal of Electronics and Information Technology,2018, 40(12):2812 - 2819.

[17] 阳理理. 基于人工神经网络的遥感图像车辆检测[D]. 南宁:广西大学,2018.
Yang L L. Vehicle detection based on artificial neural network in remote sensing images[D]. Nanning:Guangxi University,2018.

[18] 孙秉义. 基于遥感图像处理的交通量检测与分析[D]. 上海:上海交通大学,2019.
Sun B Y. Traffic volume detection and analysis based on remote sensing images processing[D]. Shanghai:Shanghai Jiaotong University,2019.

[19] Xia G S,Bai X,Ding J,et al. DOTA:A large - scale dataset for object detection in aerial images[C]// 2018 IEEE Conference on Computer Vision and Pattern Recognition. IEEE,2018:3974 - 3983.

[20] 黄国捷. 基于深度学习的遥感图像车辆目标检测[D]. 苏州:苏州大学,2019.
Huang G J. Vehicle target detection from remote sensing images based on deep learning[D]. Suzhou:Soochow University,2019.

[21] Ji H,Gao Z,Mei T,et al. Improved Faster R - CNN with multiscale feature fusion and homography augmentation for vehicle detection in remote sensing images[J]. IEEE Geoscience and Remote Sensing Letters,2019,16(11):1761 - 1765.

[22] Rottensteiner F,Sohn G,Jung J,et al. The ISPRS benchmark on urban object classification and 3D building reconstruction[J]. ISPRS Annals of Photogrammetry,Remote Sensing and Spatial Information Sciences,2012,1 - 3:293 - 298.

[23] Razakarivony S,Jurie F. Vehicle detection in aerial imagery:A small target detection benchmark[J]. Journal of Visual Communication and Image Representation,2016,34:187 - 203.

[24] Long J,Shelhamer E,Darrell T. Fully convolutional networks for

semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,39(4):640 – 651.

[25] 梁哲恒,黎 宵,邓 鹏,等. 融和多尺度特征注意力的融合遥感影像变化检测方法[J]. 测绘学报,2022,51(5):668 – 676.
Liang Z H,Li X,Deng P,et al. Remote sensing images change detection fusion method integrating multi – scale feature attention [J]. Acta Geodaetica et Cartographica Sinica,2022,51(5):668 – 676.

[26] Yang X,Sun H,Sun X,et al. Position detection and direction prediction for arbitrary – oriented ships via multitask rotation region convolutional neural network[J]. IEEE Access,2018,6:50839 – 50849.

[27] Fu Y,Wu F,Zhao J. Context – Aware and depth wise – based detection on orbit for remote sensing image[C]//2018 24th International Conference on Pattern Recognition(ICPR). IEEE,2018:1725 – 1730.

[28] Li Q,Mou L,Xu Q,et al. $R^3$ – Net:A deep network for multioriented vehicle detection in aerial images and videos[J]. IEEE Transactions on Geoscience and Remote Sensing,2019,57(7):5028 – 5042.

[29] Zhang Z H,Guo W W,Zhu S N,et al. Toward arbitrary – oriented ship detection with rotated region proposal and discrimination networks[J]. IEEE Geoscience and Remote Sensing Letters,2018,15(11):1745 – 1749.

[30] 林 钊. 基于深度学习的遥感图像舰船目标检测与识别[D]. 长沙:国防科技大学,2018.
Lin Z. Ship detection and recognition in remote sensing images based on deep learning[D]. Changsha:National University of Defense Technology,2018.

[31] 刘万军,高健康,曲海成,等. 多尺度特征增强的遥感图像舰船目标检测[J]. 自然资源遥感,2021,33(3):97 – 106. doi:10.6046/zrzyyg.2020372.
Liu W J,Gao J K,Qu H C,et al. Ship detection based on multi – scale feature enhancement of remote sensing images[J]. Remote Sensing for Natural Resources,2021,33(3):97 – 106. doi:10.6046/zrzyyg.2020372.

[32] 许德刚,王 露,李 凡. 深度学习的典型目标检测算法研究综述[J]. 计算机工程与应用,2021,57(8):10 – 25.
Xu D G,Wang L,Li F. A review of typical object detection algorithms based on deep learning[J]. Computer Engineering and Applications,2021,57(8):10 – 25.

[33] 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. IEEE,2016:779 – 788.

[34] RedmonJ,Farhadi A. YOLO9000:Better,faster,stronger[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE,2017:6517 – 6525.

[35] Redmon J,Farhadi A. YOLOv3:An incremental improvement[EB/OL]. (2019 – 12 – 25)[2021 – 12 – 29]. https://arxiv.org/abs/1804.02767.

[36] Bochkovskiy A,Wang C Y,Liao H. YOLOv4:Optimal speed and accuracy of object detection[EB/OL]. (2020 – 04 – 23)[2021 – 12/29]. https://arxiv.org/abs/2004.10934.

[37] Liu W,Anguelov D,Erhan D,et al. SSD:Single shot multibox detector[J]. Springer,Cham,2016,9905:21 – 37.

[38] Fu C,Liu W,Ranga A,et al. DSSD:Deconvolutional single shot detector[EB/OL]. (2017 – 01 – 23)[2021 – 12/29]. https://arxiv.org/abs/1701.06659v1.

[39] Tian Z,Shen C H,Chen H,et al. FCOS:Fully convolutional one – stage object detection[C]// 2019 IEEE/CVF International Conference on Computer Vision(ICCV). IEEE,2020:9626 – 9635.

[40] 李圣琀,邵峰晶. 基于深度学习的轻量遥感图像车辆检测模型[J]. 工业控制计算机,2020,33(6):66 – 69.
Li S H,Shao F J. Vehicle detection model of light weight remote sensing image based on deep learning[J]. Industrial Control Computer,2020,33(6):66 – 69.

[41] Etten A V. You only look twice:Rapid multi – scale object detection in satellite imagery[EB/OL]. (2018 – 05 – 24)[2021 – 12 – 29]. https://arxiv.org/abs/1805.09512.

[42] 彭新月,张吴明,钟若飞. 改进 YOLOv3 模型的 GF – 2 卫星影像车辆检测[J]. 测绘科学,2021,46(12):147 – 154.
Peng X Y,Zhang W M,Zhong R F. GF – 2 satellite image vehicle detection based on improved YOLOv3 model[J]. Science of Surveying and Mapping,2021,46(12):147 – 154.

[43] 汤田玉. 基于深度学习的高分辨率光学遥感影像车辆目标检测方法研究[D]. 长沙:国防科技大学,2017.
Tang T Y. Deep convolutional neural network based vehicle detection methods on high resolution optical remote sensing images[D]. Changsha:National University of Defense Technology,2017.

[44] 侯 涛,蒋 瑜. 改进 YOLOv4 在遥感飞机目标检测中的应用研究[J]. 计算机工程与应用,2021,57(12):224 – 230.
Hou T,Jiang Y. Resrarch of improved YOLOv4 in remote sensing aircraft target detection[J]. Computer Engineering and Applications,2021,57(12):224 – 230.

[45] 赵鹏飞,谢林柏,彭 力. 融合注意力机制的深层次小目标检测算法[J]. 计算机科学与探索,2022,16(4):927 – 937.
Zhao P F,Xie L B,Peng L. A deep small target detection algorithm based on attention mechanism[J]. Computer Science and Technology,2022,16(4):927 – 937.

[46] 王明阳,王江涛,刘 琛. 基于关键点的遥感图像旋转目标检测[J]. 电子测量与仪器学报,2021,35(6):102 – 108.
Wang M Y,Wang J T,Liu C. Detection of rotating targets in remote sensing images based on key points[J]. Journal of Electronic Measurement and Instrument,2021,35(6):102 – 108.

[47] 唐建宇,唐春晖. 基于旋转框和注意力机制的遥感图像目标检测算法[J]. 电子测量技术,2021,44(13):114 – 120.
Tang J Y,Tang C H. Remote sensing image target detection algorithm based on rotating frame and attention mechanism[J]. Electronic Measurement Technique,2021,44(13):114 – 120.

[48] 陈 俊. 基于 R – YOLO 的多源遥感图像海面目标融合检测算法研究[D]. 武汉:华中科技大学,2019.
Chen J. Research on maritime target fusion detection in multi – source remote sensing images based on R – YOLO[D]. Wuhan:Huazhong University of Science and Technology,2019.

[49] 谢俊章,彭 辉,唐健峰,等. 改进 YOLOv4 的密集遥感目标检测[J]. 计算机工程与应用,2021,57(22):247 – 256.
Xie J Z,Peng H,Tang J F,et al. Improved YOLOv4 for dense remote sensing target detection[J]. Computer Engineering and Applications,2021,57(22):247 – 256.

[50] 杨治佩,丁 胜,张 莉,等. 无锚点的遥感图像任意角度密集目标检测方法[J]. 计算机应用,2022,42(6):1965 – 1971.
Yang Z P,Ding S,Zhang L,et al. An arbitrary angle dense target detection method for remote sensing images without anchor points [J]. Computer Application,2022,42(6):1965 – 1971.

[51] 张宏群,班勇苗,郭玲玲,等. 基于 YOLOv5 的遥感图像舰船的检测方法[J]. 电子测量技术,2021,44(8):87 – 92.
Zhang H Q,Ban Y M,Guo L L,et al. Remote sensing images ship detection method based on YOLOv5 [J]. Electronic Measurement Technology,2021,44(8):87 – 92.

[52] 张玉莲. 光学图像海面舰船目标智能检测与识别方法研究[D]. 长春:中国科学院大学(中国科学院长春光学精密机械与物理研究所),2021.
Zhang Y L. Research on intelligent detection and recognition methods of ship targets on the sea surface in optical images[D]. Changchun:University of Chinese Academy of Sciences (Changchun Institute of Optics,Fine Mechanics and Physics,CAS),2021.

[53] Li K,Wan G,Cheng G,et al. Object detection in optical remote sensing images:A survey and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2020,159:296 – 307.

# A review and adaptability study of deep learning models for vehicle detection based on high – resolution remote sensing images

LYU Yanan[1], ZHU Hong[1], MENG Jian[2], CUI Chengling[3],SONG Qiqi[1]

(1. *College of Ecology Environment*, *Institute of Disaster Prevention*, *Langfang* 065201, *China*; 2. *College of Earth Sciences*, *Institute of Disaster Prevention*, *Langfang* 065201, *China*; 3. *Beijing Geoway Information – Technology Co.*, *Ltd.*, *Beijing* 100043, *China*)

**Abstract**: Vehicle detection is a hot research topic in the fields of computer vision, photogrammetry, and remote sensing. With the continuous development of deep learning technology, vehicle detection based on remote sensing images has been applied in fields such as smart city construction and intelligent transportation. This study systematically summarized existent vehicle detection algorithms based on remote sensing images and deep learning models and highlighted the classification, analysis, and comparison of one – stage and two – stage vehicle detection algorithms. Moreover, this study summarized the key technologies of vehicle detection in large – scale and complex backgrounds and analyzed the advantages and disadvantages of mainstream deep learning models of vehicle detection based on remote sensing images. Experiments were conducted to evaluate the YOLOv5, Faster – RCNN, FCOS, and SSD algorithms using DOTA and DIOR datasets. The vehicle detection precision based on the DOTA dataset was 0.695, 0.410, 0.370, and 0.251, respectively and that based on the DIOR dataset was 0.566, 0.243, 0.231, and 0.154, respectively. The experimental results show that the small target scale is still the main factor restricting the vehicle detection performance based on remote sensing images and that the application of deep learning models to the detection of small targets is to be further improved. Finally, based on public datasets and the analysis of existing algorithms, this study proposed the solution and development trend of vehicle detection based on remote sensing images in large – scale and complex backgrounds.

**Keywords**: remote sensing image; vehicle detection; deep learning; analysis method

(责任编辑: 张 仙)