

文章编号: 1672-8785(2022)05-0041-08

基于深度学习的无人机遥感 影像车辆检测

谭路文 哈斯巴干* 陈超民 谢 璇

(上海师范大学环境与地理科学学院, 上海 200234)

摘 要: 伴随着城市的发展, 车辆数量在不断地增加。这一现象不仅增加了城市拥挤状态, 而且还促使交通事故频发。要提高城市治理能力, 就必须提高对城市车辆的监测能力。使用无人机对上海、赤峰地区的四个场景进行了低空摄影, 获取了航空遥感影像数据, 然后结合深度学习的 Unet 卷积神经网络技术对无人机影像中的车辆进行了单目标提取。结果表明, 深度学习对无人机影像中车辆的识别能力远高于传统机器学习中的随机森林方法, 达到了 99% 的超高精确度, 且每个场景内汽车数的估算结果与真实数量极其接近。根据研究结果可知, 将无人机和深度学习技术相结合的车辆检测方法具备实时性和现实可行性, 可为城市的车辆实时监测和交通管理提供可靠的技术手段。

关键词: 深度学习; 无人机; 遥感; 车辆检测; 城市管理

中图分类号: TP751 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2022.05.007

Vehicle Detection from UAV Remote Sensing Images Based on Deep Learning

TAN Lu-wen, HASI Ba-gan*, CHENG Chao-min, XIE Xuan

(School of Environmental and Geographical Sciences, Shanghai Normal University, Shanghai 200234, China)

Abstract: With the development of the city, the number of vehicles is increasing. This phenomenon not only increases urban congestion, but also leads to frequent traffic accidents. In order to improve the ability of urban governance, it is necessary to improve the monitoring ability of urban vehicles. In this paper, UAV is used to take low-altitude photography of four scenes in Shanghai and Chifeng area, and aerial remote sensing image data are obtained. Then, single target extraction is carried out for vehicles in UAV images combined with Unet convolutional neural network technology of deep learning. The results show that the ability of deep learning to recognize vehicles in UAV images is much higher than the random forest method in traditional machine learning, which reaches an ultra-high accuracy of 99%. And the estimation result of the number of cars in each scene is very close to the real number. According to the research results, the vehicle detection method combining UAV and deep learning technology has real time and practical feasibility, which can provide reliable technical means for real-time vehicle monitoring and traffic management in cities.

收稿日期: 2021-12-16

基金项目: 上海市科技计划项目(22010503600); 国家自然科学基金项目(41771372)

作者简介: 谭路文(1996-), 男, 江西赣州人, 硕士研究生, 主要从事无人机遥感与深度学习方面的研究。

***通讯作者:** E-mail: hasibagan@staff.shnu.edu.cn

Key words: deep learning; UAV; remote sensing; vehicle detection; urban management

0 引言

随着社会科技的发展,汽车已经成为人们日常出行的必要交通工具。车辆的监控成为了当今社会的一个大难题。快速、准确地获取城市内部的车辆时空分布信息,将帮助城市管理人员更好地进行城市治理^[1-3]。卫星遥感能够在大气尺度上监视整个地球的动态变化,但对于汽车这类小目标物体的检测,它在空间与时间分辨率上都难以满足要求^[4]。而大量安装固定摄像头虽然可以对地面车辆进行长期监控,但其缺点也很明显,比如检测场景有限、安装成本高以及存在较多的监控死角。无人机遥感具有快捷、成本低廉、易于操作等优点,能很好地弥补卫星遥感和固定摄像头的缺陷^[5-6]。

近年来,机器学习技术的出现为遥感影像解译提供了强大的技术手段。但是传统的浅层机器学习方法需要根据任务目的去人工筛选特征,图像解译的质量很容易受光照、阴影、图像质量等因素的影响。随着机器学习的深入发展,科学家们又提出了深度学习的方法。它是一种具有多个隐藏层和更深组合的人工神经网络,在目标检测方面表现出良好的性能^[7-8]。

鉴于此,本文采用无人机对上海、赤峰的四个车辆密集的场景进行了低空摄影,以获取

城市的航空遥感影像数据集。对获取到的影像进行融合及图像增强处理,再利用深度学习的Unet 卷积神经网络技术对遥感影像进行训练,可获得高精度的模型。利用模型基本能够预测出场景中的每一个车辆。通过对车辆结果进行统计计算,还能够较为准确地预测出场景内的车辆数量。本文系统性地论述了从无人机数据的前期获取到中期处理,再到后期深度学习的训练、预测的每个步骤。结果表明,该方法具备现实可行性和实时性,可为城市交通和社区管理提供一种极具价值的技术手段。

1 研究数据

1.1 研究区域

本文的主要研究区域为上海市以及内蒙古自治区赤峰市。上海市的总面积为 6340.5 km²。作为中国第一大城市,上海人口高度聚集且车辆众多;常住人口有 2487.09 万。本文中车辆检测的道路、车站和小区三个场景均在上海市。

赤峰市的总面积为 90021 km²,常住人口为 403.5967 万;地域辽阔,人口密度低,停车场多为室外露天停车场。而上海市许多的停车场设置在地下和建筑物里面。因此,本文将停车场研究区域选址在内蒙古赤峰市。上海市和赤峰市的研究区域位置如图 1 所示。

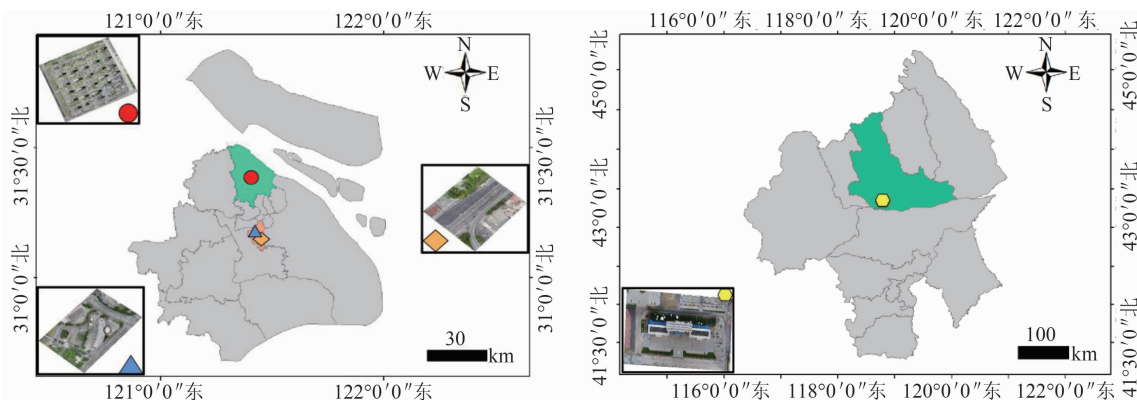
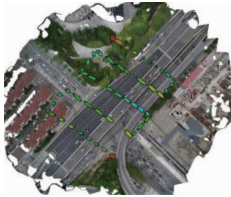
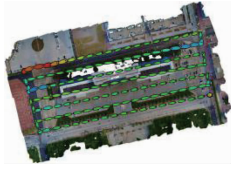


图 1 研究区域位置图(左为上海市,右为赤峰市)

表 1 飞行路线和参数表

航线规划	飞行高度	飞行速度	拍照间隔	分辨率	照片数量
	100 m	40 km/h	1 s	2.56 cm/pix	52
	100 m	40 km/h	2 s	2.58 cm/pix	111
	100 m	30 km/h	2 s	2.64 cm/pix	236
	60 m	30 km/h	2 s	1.59 cm/pix	201

1.2 数据获取

本文的实验数据均使用大疆商业无人机采集, 其型号为精灵 phantom4 Pro 暗夜版(见图 2)^[9]。该无人机的云台相机能够保障其在飞行晃动过程中仍然可以获取高清影像^[10]。

无人机操控及航线规划软件为 Litchi 4.17.0。对于四个不同的场景, 我们设定了四

条不同的飞行轨道, 总共获取 600 张影像。其中部分区域重叠率过高, 标注意义不大, 因此舍弃部分图像。最终用来制作样本标签的图像共有 436 张^[11]。表 1 列出了四个场景的无人机飞行路线及参数^[12]。

2 实验方法

2.1 影像拼接融合

由于无人机飞行高度的限制, 每次摄影成像的范围只有一小部分。若要获取一幅完整的研究区场景, 则需对每条飞行轨迹上获取的照片进行拼接融合, 最终生成 RGB 三波段的正射镶嵌影像以及包含地物高程信息的 DSM 影像。

在本实验的四个场景中, 影像的航向重叠率大于 75%, 旁向重叠率大于 45%, 图像拼接过程的关键步骤就是寻找影像之间的特征点, 进而寻求影像重叠区域的同名像点,



图 2 大疆精灵 4 Pro 机型图

最终实现相邻影像集合坐标调整。本文使用高斯差分尺度空间 DOG 来检测极值点。高斯差分尺度空间函数如下：

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma)] \otimes I(x, y) \quad (1)$$

式中, $I(x, y)$ 为图像数据, k 为相邻尺度空间的比例因子, $G(x, y, \sigma)$ 为尺度可变高斯函数。

二维高斯函数 $G(x, y, \sigma)$ 定义为

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

式中, σ 为尺度空间因子, 是高斯正态分布的方差, 表明原始图像被高斯模糊的次数。其值越小, 图像清晰程度越高; 反之, 图像清晰程度越低^[13]。

2.2 图像增强

本文基于 Python 的 Opencv 库对无人机影像进行图像增强, 进而使试验获得更好的效果。众所周知, 在一天中的不同时段去获取遥感影像会有不同的亮度、色差、阴影和噪声。在本文试验中, 通过变亮处理、变暗处理、椒盐噪声处理、模糊处理等操作来实现图像增强, 从而增加深度学习模型的鲁棒性^[14-15]。图像增强样例如图 3 所示。

2.3 Unet 卷积神经网络

Unet 网络结构由 Ronneberger O 等人在

2015 年 ISBI 竞赛中提出, 即用收缩通道和扩张通道构成一个 U 型结构, 因此命名为 U-Net^[16]。

收缩通道的作用是提取影像的深层特征。先进行两次重复的 3×3 卷积, 对每次卷积的结果使用 ReLU 函数进行激活, 再通过 2×2 的滤波对特征图进行最大池化操作, 然后重复上述步骤不断提取影像中更深层次的特征。

扩张通道的作用是还原图像的特征信息。先经过一次 2×2 的上采样操作, 然后将上采样后的特征图与收缩通道处对应的特征图进行连接操作(得到的特征图中同时包含图像的浅层特征和深层特征信息), 接着进行两次重复的 3×3 卷积操作, 将特征图的通道数减少一半。

在最后一层中使用 1×1 的卷积, 将 64 个特征映射到所要分类的 2 个类别中。图像中的每个像元都分别归类到其中一类中。Unet 网络架构如图 4 所示。

3 实验结果

3.1 结果分析

在深度学习训练过程中, 为减轻电脑运算量, 将原始图像的尺寸由 5472×3078 重采样为 1280×768 。在图像输入过程中, batch_size 大小设置为 1。初始学习率设置为 10^{-3} 以加快



图 3 图像增强处理

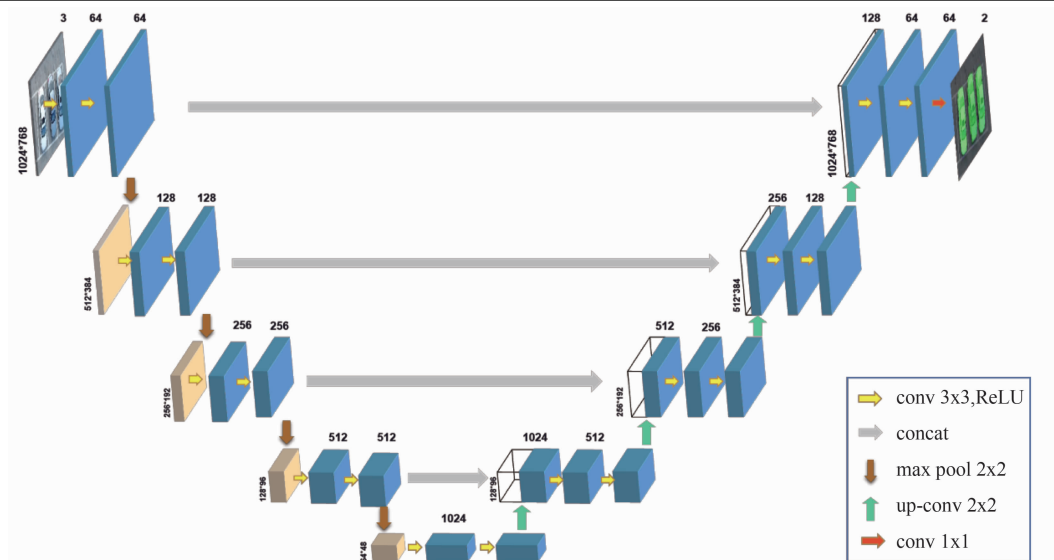


图 4 U-Net 网络结构模型

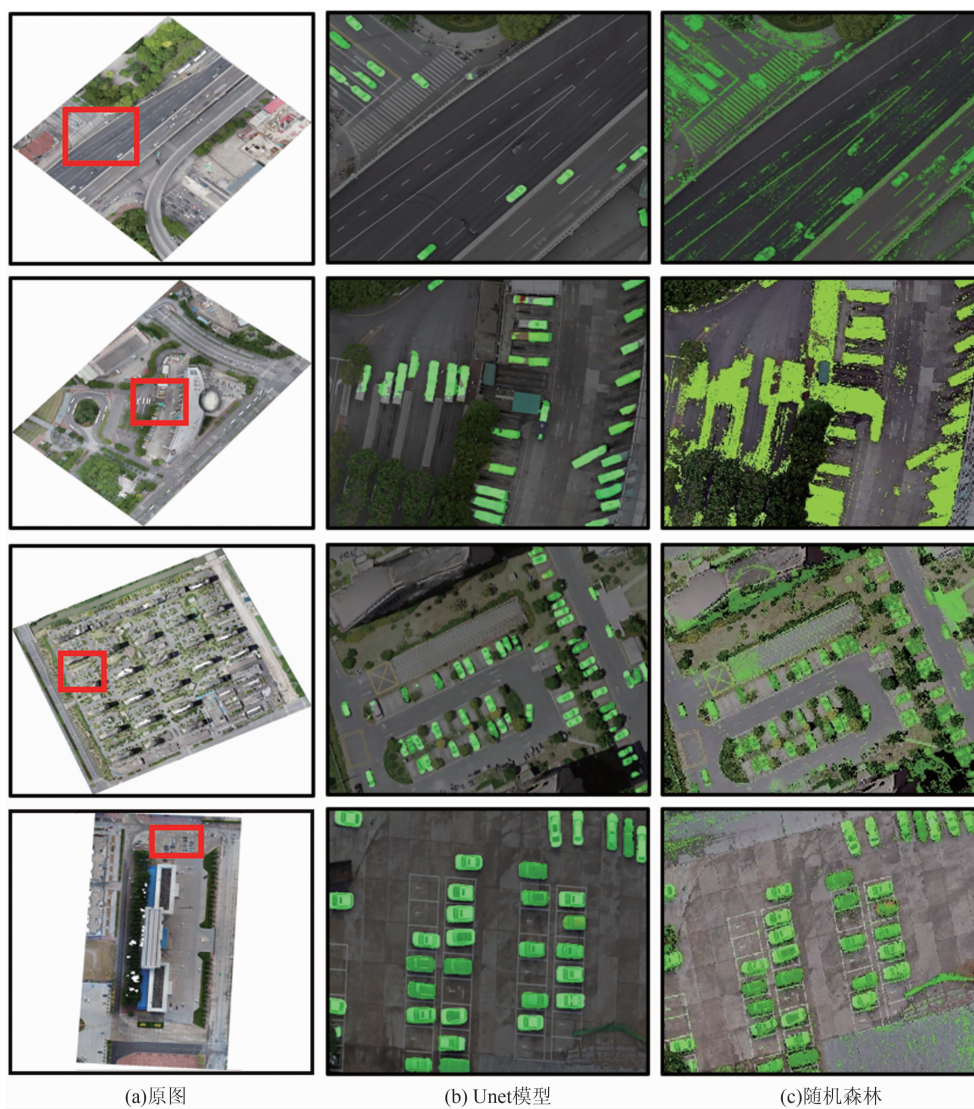


图 5 实验结果图

表 2 车量估算表

	车辆类型	飞行高度	车辆尺寸 S_{avg}	预估车量	实际车量
场景一	小轿车	60 m	34000	93	95
场景二	小轿车	100 m	12000	853	897
场景三	小轿车、大巴车	100 m	22000	202	213
场景四	小轿车	100 m	12000	84	85

收敛速度。当模型精度趋于稳定时,学习率改为 10^{-4} ,使模型能够达到更高的精度。

对照试验使用的是传统机器学习的随机森林算法。随机森林主要设置树数和树深两个超参数。通过多次试验可知,当树数设置为 1000 且树深设置为 30 时,效果最佳。随机森林的参数数量远比深度学习模型少,无法很好地提取出目标车辆的众多特征信息,容易产生误分和错分的情况。因此,我们需要在数据输入时尽可能地输入更多特征的图像,以保证车辆检测的精度^[17-18]。在对照试验中,除了输入原始影像的 RGB 三波段外,还增加了 DSM 波段,使得分类器能够通过地物的高程信息来区分车辆和背景。两种不同方法对四个场景的车辆检测实验结果如图 5 所示。

从图 5 中可以很明显地看出,基于深度学习的无人机遥感影像分类结果基本能把四个场景中的每一辆车都准确识别出来,在场景二和场景三中有少部分车辆识别不完全。而随机森林方法在车辆识别中的效果较差且错分情况较多,容易将植被、路面和低矮候车棚等背景信息错分成车辆。

由于深度学习的方法对车辆检测效果好,我们不仅可以得出地面车辆的空间分布图,还能够通过统计分析得出汽车的大致数量^[19-21]:

$$N_i = \text{round}\left(\frac{A_i}{S_{avg}}\right), i = 1, \dots, R_n \quad (3)$$

式中, R_n 为检测场景总数, A_i 为每个场景的车辆检测面积, S_{avg} 为平均车辆尺寸(即平均车辆像元数), N_i 为该场景的车辆估计数。由于四个场景中的无人机飞行高度不

同,车辆类型也不同,所以每个场景中 S_{avg} 的取值也不同。

表 2 为具体车量估算表。可以看出,车辆的预估数量与实际数量较为接近。通过统计计算能够大致估算出某个场景中的车辆数量情况。同时,预估车辆数都要比实际车辆数少。造成该现象的原因是在无人机影像中,少部分车辆位于树荫下面,检测时只有部分车身能被识别出来。

3.2 精度评定

通常我们使用准确率和交叉熵损失函数来评估深度学习模型的性能^[22]。交叉熵损失函数在二分的情况下,模型最后需要预测的结果只有两种。对于两个类别,我们预测得到的概率为 p 和 $1-p$,其表达式为

$$C = -\frac{1}{n} \sum_i [y_i \cdot \log(p_i) + (1 - y_i) \times \log(1 - p_i)] \quad (4)$$

式中, n 表示总的样本数量; y_i 表示样本 i 的 label, 正类为 1, 负类为 0; p_i 表示样本 i 预测为正类的概率。

通俗来说,准确率 A_c 是指模型预测正确的结果所占的比例:

$$A_c = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

在获得预测结果图后,分别计算结果图像中 TP 、 FP 、 TN 和 FN 像素的数量。 TP 表示预测为正样本像元,实际也为正样本像元; FP 表示预测为正样本像元,实际为负样本像元; TN 表示预测为负样本像元,实际也为负样本像元; FN 表示预测为负样本像元,实际为正样本像元。表 3 列出了 Unet 模型的准确率及损失率。

表 3 Unet 模型精度评价表

	准确率	损失率
训练集	99.88%	0.3%
验证集	99.43%	2.62%

本文试验中共有 2180 张样本图片。按 6:4 的比例将它们分成训练集和验证集。其中, 训练集中共有 1308 张样本图片, 验证集中共有 872 张样本图片。总共训练 35 个 epochs。结果表明, Unet 车辆目标检测模型在训练集中的准确率为 99.88%, 交叉熵损失值为 0.3%; 在验证集中的准确率为 99.43%, 交叉熵损失值为 2.62%。无论从数值还是结果图像中都可以看出, 该模型在车辆检测方面的性能表现非常好。

4 讨论与展望

为了助力美好城市建设目标并提高社会管理人员对城市车辆的监测能力, 本文采用深度学习方法来实现无人机遥感影像中的车辆目标检测。与传统机器学习方法对比后发现, 深度学习方法是目前能够实现车辆目标信息提取的最佳选择, 其车辆目标检测能达到 99% 的超高精确度。本文详细阐述了从无人机航线规划到图像增强再到模型训练预测的完整过程, 证明该方法具有现实可行性。无人机简易、便捷的特点也使社会管理人员能够随时随地利用无人机获取到最新的航空遥感影像。将它们再制成相应的车辆空间分布图并预估出场景中的车辆数, 从而辅助城市的治理。

深度学习方法在车辆检测中的精度已经达到极高的地步, 在未来的无人机车辆检测研究中继续提高检测精度的意义已经不大。我们对无人机遥感的车辆检测的研究方向将是细分检测目标, 比如将车辆细分成小轿车、公交车(大巴车)、货车、电动车等。由于不同类型的车辆对城市拥堵状况的影响情况不一样, 细分车辆类型将会进一步改善城市生活。

参考文献

[1] Ammar A, Koubaa A, Ahmed M, et al. Vehicle

Detection from Aerial Images Using Deep Learning: A Comparative Study [J]. *Electronics*, 2021, **10**(7): 820.

[2] 姜尚洁, 罗斌, 刘军, 等. 基于无人机的车辆目标实时检测 [J]. *测绘通报*, 2017, **63**(S1): 164-168.

[3] Kumar S, Rajan E G, Rani S, et al. A Study on Vehicle Detection through Aerial Images: Various Challenges, Issues and Applications [C]. Greater Noida: International Conference on Computing, Communication, and Intelligent Systems, 2021.

[4] 杨文哲. 基于深度学习的无人机小目标车辆检测方法 [D]. 西安: 西安电子科技大学, 2020.

[5] Osco L P, Jose M J, Ana P M R, et al. A Review on Deep Learning in UAV Remote Sensing [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2021, **102**: 102456.

[6] 项阳. 基于深度学习的无人机影像车辆识别[J]. *信息与电脑(理论版)*, 2018, **25**(9): 71-73.

[7] 赵爽, 黄怀玉, 胡一鸣, 等. 基于深度学习的无人机航拍车辆检测 [J]. *计算机应用*, 2019, **39**(S2): 91-96.

[8] 程擎, 范满, 李彦冬, 等. 无人机航拍图像语义分割研究综述 [J]. *计算机工程与应用*, 2021, **57**(19): 57-69.

[9] 路佳, 杨立树, 姚益峰. Phantom 4 Pro 无人机倾斜摄影测量技术 [J]. *世界有色金属*, 2021, **28**(2): 180-181.

[10] Guo Y, Chen S, Wu Z, et al. Integrating Spectral and Textural Information for Monitoring the Growth of Pear Trees Using Optical Images from the UAV Platform [J]. *Remote Sensing*, 2021, **13**(9): 1795.

[11] Inoue Y, Masaki Y. Drone-based Remote Sensing of Crops and Soils and Its Application to Smart Agriculture [J]. *Journal of The Remote Sensing Society of Japan*, 2017, **37**(3): 224-235.

[12] 朱强. 多旋翼无人机航线优化及航测精度分析 [D]. 郑州: 郑州大学, 2016.

[13] 贾银江. 无人机遥感图像拼接关键技术研究 [D]. 哈尔滨: 东北农业大学, 2016.

[14] 王浩, 张叶, 沈宏海, 等. 图像增强算法综述 [J]. *中国光学*, 2017, **10**(4): 438-448.

- [15] 孙晓凯, 倪卿元, 陈文强. 图像增强方法在深度学习图像识别场景应用中的可行性研究 [J]. *电信科学*, 2020, **36**(S1): 172–179.
- [16] Ronneberger O, Philipp F, Thomas B. U-net: Convolutional Networks for Biomedical Image Segmentation [C]. Cham: International conference on medical image computing and computer-assisted intervention, 2015.
- [17] 雷震. 随机森林及其在遥感影像处理中应用研究 [D]. 上海: 上海交通大学, 2012.
- [18] 姬晓飞, 石宇辰. 多分类器融合的光学遥感图像目标识别算法 [J]. *计算机技术与发展*, 2019, **29**(11): 52–56.
- [19] Ammour N, Alhichri H, Bazi Y, et al. Deep Learning Approach for Car Detection in UAV Imagery [J]. *Remote Sensing*, 2017, **9**(4): 312.
- [20] Moranduzzo T, Farid M. Automatic Car Counting Method for Unmanned Aerial Vehicle Images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, **52**(3): 1635–1647.
- [21] 蒋妮, 周海洋, 余飞鸿. 基于计算机视觉的目标计数方法综述 [J]. *激光与光电子学进展*, 2021, **58**(14): 43–59.
- [22] 周南南. 基于深度卷积神经网络的图像语义分割方法研究 [D]. 秦皇岛: 燕山大学, 2019.