

基于深度学习的遥感影像松材线虫病树提取

吴思琪

长江大学地球科学学院, 湖北 武汉
Email: 2575763177@qq.com

收稿日期: 2021年4月24日; 录用日期: 2021年5月19日; 发布日期: 2021年5月26日

摘要

松材线虫病对我国松树类物种具有极大的伤害, 需要对病虫区域进行准确高效的确定以提早防治。这种病在传播方式上具有跳跃特性。它具有传播途径多种多样、发病部位较隐蔽难以发现、病情潜伏时间长、发病速度迅速、治理不方便等特点, 严重时会导致大量松树病死, 导致环境和森林景观的严重损坏, 并可能导致严重的经济和环境损失。本论文通过使用深度学习目标检测中的RetinaNet方法, 将无人机拍摄的影像作为训练样本, 充分利用深度学习目标检测方法的优点, 将其对比SSD和YOLO v3方法的识别效果, 实现病虫树木的高效判别。对松材线虫病树区域展开定位研究, 在节省人工成本的同时能迅速防止病虫害对松树的疾病扩散, 为清除和防治病害区域扩散至更大范围提供有效帮助。

关键词

松材线虫病, 深度学习, 目标检测

Remote Sensing Image Extraction of Pine Wood Nematode Disease Tree Based on Deep Learning

Siqi Wu

School of Geosciences, Yangtze University, Wuhan Hubei
Email: 2575763177@qq.com

Received: Apr. 24th, 2021; accepted: May 19th, 2021; published: May 26th, 2021

Abstract

Pine wood nematode disease is very harmful to pine species in China, so it is necessary to determine the pest area accurately and efficiently in order to prevent it in advance. The disease has a

leaping characteristic in its mode of transmission. It has the characteristics of diversity, strong concealment of transmission route, long incubation period of disease, fast transmission speed and inconvenient management. When it is serious, a large number of pine trees will die, causing serious damage to environment and forest landscape, and may lead to serious economic and environmental damage. In this paper, through the use of RetinaNet method in deep learning target detection, the unmanned aerial vehicle images are taken as training samples, making full use of the advantages of deep learning target detection method, the recognition effect of SSD and Yolo V3 method is compared, and the efficient identification of pest trees is realized. The research on the location of pine wood nematode disease tree area can save the labor cost and prevent the spread of diseases and insect pests on pine trees quickly, which can provide effective help for clearing and controlling the spread of disease area to a wider range.

Keywords

Pine Wood Nematode Disease, Deep Learning, Target Detection

Copyright © 2021 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

松材线虫是松树的毁灭性病害,在1982年被人们发现时便迅速蔓延,对大片的松林构成了毁灭性威胁。它具有传播途径多种多样、发病部位较隐蔽难以发现、病情潜伏时间长难以提前察觉、发病速度迅速、治理不方便等特点[1]。树木针叶颜色逐渐变成黄色或者褐色是被感染松树的表面症状,而红褐色的针叶是重病时期松树。为了防治松材线虫对森林资源的伤害,必须及时清理死树或者染上疾病的松树[2]。此外,一旦病情蔓延会造成大量的松树死亡,如果不及早发现和治疗,松材线虫病将出现将会给国家带来巨大的经济损失以及对森林的生态产生一定的破坏。根据不完全统计,松材线虫病已在浙江、广东、四川的244个县级行政区域及其他16个省出现,累计死亡松树达到500多万株,造成的直接和间接经济损失达到约275亿元[3]。

如今病虫区域的检测有很多种,其中人工方式监测病害树木虽然可行但部分发病区耗时耗力、成本高且人力难以到达。但是如果可以利用将无人机技术、卫星导航定位技术、计算机技术等手段相结合的方式[4],便能对松材线虫病树进行实时的监测,这样有害生物监测和预警就会变得高效有用,还可以解决一旦病情爆发发现不及时的重大问题[5]。

目前,利用遥感技术识别枯木的方法多种多样[6]。在对研究区马尾松进行多光谱校正和分割后,使用依赖于先验知识的决策树分类方法对马尾松进行信息提取[7]。并且在监测过程中,还可以通过松树的纹理和颜色的特征,利用隶属函数法对患病松树进行识别,进而可以将松树的疾病程度分为轻度、中度和重度。利用计算机图像分析技术比如说图像分割算法提取无人机航拍图像中的害虫区域。利用面向对象分类算法和支持向量机、k近邻、随机森林等机器学习分类算法对无人机高光谱图像中的植被种类和灾害等级进行分类[8]。

与传统的目标探测相比,深度学习模型具有较强的表现能力,多尺度问题上的特征提取和时间效率上有优势。这种病虫树的有效提取,不仅可以降低人工筛选的成本,而且可以对人迹罕至的深山林区进行监测,提高病虫树的筛选效率,及时发现病虫树和枯树,防止进一步扩大病虫害的范围,并减少因病

虫害造成的经济损失[9]。本文主要研究深度学习目标检测的方法 RetinaNet 结合 ResNet 主干网络对松材线虫病枯死树训练样本进行训练,训练出来的模型能很好地识别受灾地区松材线虫病树的位置。

2. 基于深度学习的目标检测发展

在过去十年中,自然图像对象检测算法可分为两个不同的阶段,第一阶段是在 2013 年的期间基于传统的人工时代,第二个阶段则是以深度学习为基础进行的目标检测。如果从目标检测技术演变的角度来看,技术的发展则分别经历了包围框回归时期、多参考窗口时期、神经网络兴起时期、难样本挖掘与聚焦时期和多尺度多窗口检测时期[10]。

此外,在图像识别领域,目前深度学习通过构建神经网络学习的方法使错误率降低,可以分为基于候选区域的方法和回归方法。前者是 two-stage 类,主要包括 SPP-Net [11]、RCNN [12]、Fast R-CNN [13]、Faster R-CNN [13]等,包括候选区域提取,特征提取,卷积神经网络分类和回归的边界框。这种基于候选区域的检测方法与以往基于机器学习的检测方法相比检测精度有了很大的提高,但处理速度不能满足实时检测的要求,与病树检测对时效性要求较高的要求违背。后者是基于回归的 one-stage 类方法,主要有 SSD [14]和 YOLO [15]系列算法。目标检测任务被这些算法看成回归问题,直接返回目标类别和目标边界框,大大提高了检测速度,但检测精度不及前者算法。正是因为检测速度比 two-stage 快,因此精度方面相比就有所下降。其根本原因就是 one-stage 类别虽然检测速度快但却受制于类别不平衡。类别不平衡是指在目标检测算法的前期会生成一大波的包含物体的一个紧致矩形边界框。在检测一幅常规的图片中,需要识别的对象并不会会有很多,这意味着绝大多数的先验框属于不需要识别的背景,导致先验框数量爆炸。先验框中属于背景的框远远超出识别物体的边界框,所以如果分类器会自动地把所有先验框统一归类为背景框,这样精度即使可以刷得很高,但很明显不符合要求。在第一次生成类别极不平衡的先验框中 one-stage 类型的目标检测直接就进行难度极大的细分类,希望能直接输出先验框和检测的结果。而损失函数无法从根本上解决类别的巨大不平衡,这样还是会导致模型的训练失败。因此,one-stage 类型的检测在保持住了检测速度的同时却丧失了检测精度[16]。

3. RetinaNet 的突破

3.1. Focal Loss 的特点

RetinaNet 其实在目标检测的框架上相对于 YOLO、SSD 没有什么特别改进,脸书团队通过后续实验证明了 Focal loss 可在 one-stage 类型中的成功使用,最终以 two-stage 系相似或者更加优秀的效果和更快的速率实现。

Focal loss 的特点在于两个方面一是控制了正负样本的权重,二是控制容易分类和难分类样本的权重[17]。正负样本就是一张图可能会生成巨大数量的候选框,只有少部分有我们要检测的物体,有这些物体的候选框就是正样本,没有就是负样本。对于常用的交叉熵 loss 公式(1)。

$$CE(p, y) = CE(pt) = -\log(pt) \quad (1)$$

显然需要在公式上做改动减少正样本和负样本的影响,可以通过增加系数 α_i 实现即将交叉熵变为公式(2)。

$$CE(pt) = -\alpha_i \log(pt) \quad (2)$$

当真实标签为 1 的时候, $\alpha_i = \alpha$; 当真实标签为 otherwise 也就是 0 的时候, $\alpha_i = 1 - \alpha$, α 的范围也是 0 到 1。当把 α 的值设置在 0 到 0.5 之间的时候公式(3)和公式(4)。

$$\alpha_i = \alpha (y = 1) \tag{3}$$

$$\alpha_i = 1 - \alpha (\text{otherwise}) \tag{4}$$

这两种情况中 α 的值可以看出，正样本的权重降低，而负样本的权重增加。当把 α 的值设置在 0.5 到 1 之间的时候，正样本的权重增加，而负样本的权重降低。如此一来 α 就能解决正负样本权重不平衡的问题。

同理，控制容易分类和难分类样本的权重的问题也需要得到解决。调节正负样本权重的公式(5)。

$$FL(pt) = -(1 - pt)^\gamma \log(pt) = (1 - pt)^\gamma CE(\hat{y}_i) \tag{5}$$

容易分类的样本是当输入的样本为正样本时， pt 接近于 1 意味着此时的网络可以正确识别输入进来的样本，那么权重就会很小。难分类的样本是当输入的样本为负样本时， pt 接近于 0 时，此时当前网络没有办法正确识别输入进来的样本，那么权重与容易分类的样本的权重相比会大很多。

如果 pt 的值接近于 0 时，调制系数 $(1 - pt)^\gamma$ 的值就会趋于 1，这样对交叉熵的贡献很大。如果该值接近于 1，对应的调制系数就趋于 0，也就是对于交叉熵贡献就会很小。这样 pt 能控制难以分类和容易分类样本的权重，前面提到通过控制 α 的大小能调节正负样本对 loss 的贡献。这样将这两个可控的因素合在一起就变成最终版本的 focal loss 公式(6) [18]。

$$FL(pt) = -\alpha_i (1 - pt)^\gamma \log(pt) \tag{6}$$

3.2. RetinaNet 的实现思路

由图 1 中 RetinaNet 抽象的结构图可知，主要分为三个部分。第一个部分是主干网络的特征提取，主干网络是 ResNet 深度残差网络。有一条残差边经过处理直接从某一个特征层的输入跨越到某一个特征层的输出也就是意味着将靠前若干层的某一层数据输出跨越到后面输入的数据层[19]。第二个部分是在主干特征提取网络里面加上的图像金字塔(feature pyramid net)，主要解决的是物体检测中的多尺度问题，通过简单的网络连接改变，在基本不增加原有模型计算量的情况下，大幅度提升了小物体检测的性能，它通过高层特征进行上采样和低层特征进行自顶向下的连接。第三个部分是将前面提取到的有效的特征层传输过 class + box subnets 获得预测结果。class subnet 可以用于预测每一个网格内每一个先验框的种类。box subnet 是预测每一个网格上每一个先验框调整结构。

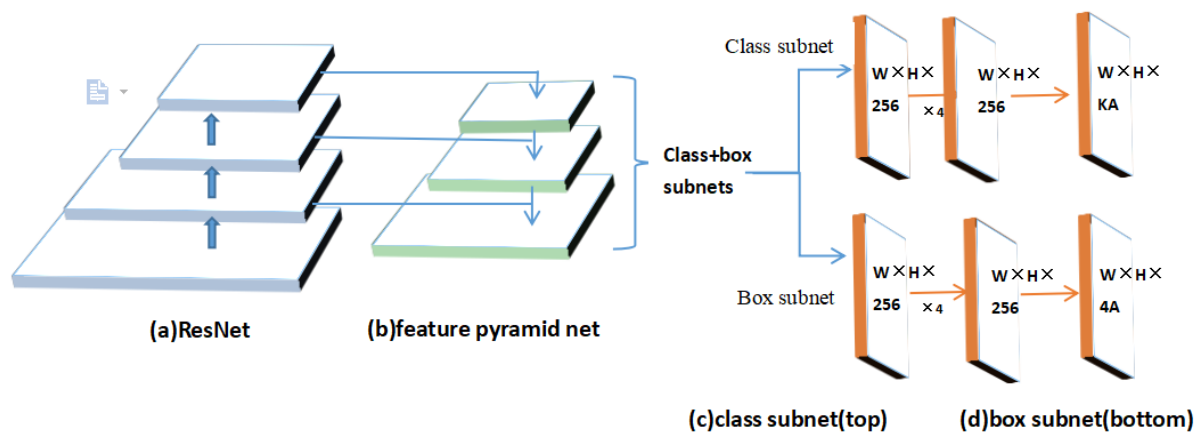


Figure 1. The network structure of RetinaNet
图 1. RetinaNet 网络结构

4. 实验过程

本次实验数据选用航空遥感数据 10 张湖北省宜昌市夷陵区航空遥感影像(大小 $13,527 \times 9113$, 分辨率 0.1 米)大图, 将其裁剪成 200 小图, 然后通过一款开源的图像标注工具 labellmg, 采集每张图片中病树的训练样本。实验选用台式机作为部署平台, 硬件信息如表 1 所示。

Table 1. Specific information of computer hardware environment

表 1. 计算机硬件环境具体信息

硬件信息			环境信息	
平台	型号	主频/内存	平台	版本
RAM	ChannelB-DIMM1	16G	tensorflow-gpu	1.13.2
CPU	i7-9700	3.0 GHz	CUDA	10.0
GPU	GTX 1650	12G	CUDNN	7.4.1.5
Windows 10	64 位		Python	3.6

病树的数据集为 VOC2007 格式, 模型的训练必须依赖数据集, 将文件夹命名为 VOCdevkit。该数据集格式中会有三类文件夹分别名为 Annotations 文件夹, ImageSets 文件夹和 JPEGImages 文件夹。Annotations 文件夹内存放的是与图像对应的检测目标的标签文件, 该文件是对图片的解释, 不同的图片对应于一个同名的 xml 文件, 也与图片文件夹里的图片相对应。ImageSets 文件夹中的 txt 文件将 VOCdevkit 里保存的图片分类为各种集合。

此外, 深度学习目标检测的模型训练过程中通过设置初始化参数、调整数据规模是提高整体训练效率的有效方式。参数的设置需要在不断调整中充分考虑到训练精度、训练速度和硬件条件。初始化参数方面, num_classes 代表需要检测的种类, 本论文只需要检测病树区域也就是病树区域, 因此 num_classes 设置为 1。验证集和数据集划分时, 设置 90% 的用于训练剩下 10% 的用于验证, 因此 val_split 设置为 0.1。训练参数方面, batch_size、learning rate 和 epochs 非常重要。batch_size 代表一次训练所选取的样本数, 它大小影响模型的优化程度和速度, 同时其直接影响到 GPU 内存的使用情况, batch_size 设置为 2。learning rate 是监督学习以及深度学习中重要的超参, 其决定着目标函数能否收敛到局部最小值以及何时收敛到最小值, 合适的学习率能够使目标函数在合适的时间内收敛到局部最小值, learning rate 设置为 0.001。epochs 为向前和向后传播中所有批次的单次训练迭代, 考虑到数据集, epochs 设置大小为 100。在本文中, 使用微软发布的 Microsoft COCO 作为预训练数据集, 该数据集包含约 30 多万张图片、200 多万个物体被标注、分为 91 个类型, 每一类物体的图片包含精确的分割信息, 是目前每张图片平均包含目标数最多的数据集。通过该数据集训练原始的 RetinaNet, 将训练后的模型参数对本文模型中相同的部分进行初始化, 然后再用自己制作的病树数据集继续训练, 这样可以实现迁移学习, 解决样本不足的问题。

5. 结果分析

结果精度评价

在目标检测模型中, 为了定量评价模型性能需要用到一些重要指标, 包括精确度(Precision)、召回率(Recall)和 F1 score。

这些评价指数依赖于 TP、FP 和 FN。TP (True Positive)就是指那些正样本被正确分类。FP (False Positives)就是指那些负样本被错误的分类为正样本。FN (False Negatives)就是正样本被错误的分类成负样本。

精确度翻译就是分类时被认为是正样本而且确实是正样本的部分占有所有分类过程中被当作是正样本的比例。那么预测为正就有两种可能了，一种就是把正类预测为正类(TP)，另一种就是把负类预测为正类(FP)，也就是精确度如公式(7)。

$$TP \div (TP + FP) \tag{7}$$

召回率公式实际上就是分类时正样本确实被正确分类的部分占所有的正样本包括被错误分类正样本的百分比。那也有两种可能，一种是把原来的正类预测成正类(TP)，另一种就是把原来的正类预测为负类(FN)召回率如公式(8)。

$$TP \div (TP + FN) \tag{8}$$

在本次实验中，在同一环境下分别使用 RetinaNet、YOLO v3 和 SSD 3 种目标检测主流 one-stage 类检测框架，并用同样的标记好的病树数据集进行训练，对最终的预测结果进行分析对比，3 种框架平台的训练识别效果如图 2 所示。其中红色框为算法检测到的病树区域，黄色框为漏检的病树区域，蓝色框为错检的病树区域。各个框架对病树的识别效果如表 2 所示。由识别效果得出，在实验区域相同的情况下基于 one-stage 类检测框架种 RetinaNet 的识别效果最好，没有错检的病树区域，少数病树区域未能被框架检测到，精确度、召回率和 F1 score 在三种框架皆为最高分别高达 100%、90.4% 和 94.9%。其次的效果是 YOLO v3，能够基本实现对病树区域的判别，但是一旦病树区域较多或者病区接近时便出现漏检和错检的情况。相比之下同样基于 one-stage 类检测框架的 SSD 的识别效果最差，不仅将区域相近的病树错检，而且还有一些明显病死的枯树没有被框架识别，但总体上看 SSD 算法和 YOLO v3 算法的精确度和召回率都很接近。

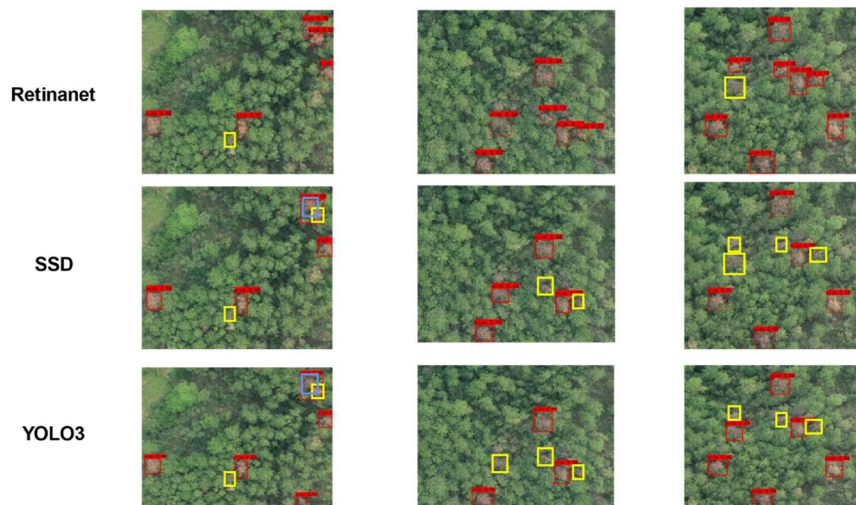


Figure 2. Multiple result graphs of disease tree region detection
图 2. 多张病树区域检测结果图

Table 2. The recognition effect of each frame on disease tree
表 2. 各个框架对病树的识别效果

检测方法	TP	FN	FP	精确度	召回率	F1 score
RetinaNet	19	2	0	100%	90.4%	94.9%
SSD	12	8	1	92.3%	60%	72.7%
YOLO V3	13	8	1	92.8%	61.9%	74.2%

6. 总结与展望

从整体上来说,实现了 RetinaNet 深度学习网络对无人机影像中的真实场景下松材线虫病树区域目标物的高效检测,比 SSD 和 YOLO v3 检测效果更好,论证了 Focal loss 的改进使 RetinaNet 有更好的优越性。实验结果显示,该方式能够有效准确地检测无人机影像中的病树区域,既能减少人工筛查的成本,更能监测人工无法到达的深山林区,提升了筛查病害树木的效率,及时发现害病树木和枯死树木,防止病害进一步扩大感染范围,减少病虫害造成的经济损失,为以后深度学习在松材线虫病树的应用提供更加先进的思路。

但同时也存在个别病树区域未识别到的现象,在后续的研究中可以通过提高病树集图片的质量、数量和多样性来给模型提供更丰富的病树特征,增强模型的识别效果;也可以通过提高计算机硬件配置、计算速度从而修改框架参数获得更好的拟合模型。

参考文献

- [1] 曾全,孙华富,杨远亮,周建华,杨超. 无人机监测松材线虫病的精度比较[J]. 四川林业科技, 2019, 40(3): 92-95, 114.
- [2] 武红敢,牟晓伟,杨清钰,王成波. 无人机遥感技术在重庆市沙坪坝区松材线虫病监测中的应用[J]. 林业资源管理, 2019(2): 109-115.
- [3] 林孝春. 松材线虫病的危害及防控措施综述[J]. 华东森林经理, 2015, 29(3): 28-30.
- [4] 马菁. 基于遥感技术的松材线虫病早期预警[D]: [硕士学位论文]. 北京: 北京林业大学, 2012.
- [5] 张裕,杨海涛. 基于深度学习的轻量化遥感图像目标检测方法[J]. 信息技术, 2019, 43(9): 163-167.
- [6] 葛宏立. 基于多源遥感信息的松材线虫病综合监测预警系统研究——两个遥感图像专题信息提取新方法[J]. 中国科技成果, 2013(21): 37-38.
- [7] 隋雪莲,张涛,曲乔新. 深度学习在遥感影像目标识别与定位中的应用研究[J]. 科技创新与应用, 2019(34): 180-181.
- [8] 张学敏. 基于支持向量数据描述的遥感图像病害松树识别研究[D]: [硕士学位论文]. 合肥: 安徽大学, 2014.
- [9] 吴琼. 基于遥感图像的松材线虫病区域检测算法研究[D]: [硕士学位论文]. 合肥: 安徽大学, 2013.
- [10] 叶钊. 目标检测技术研究进展[C]//中国计算机用户协会网络应用分会. 中国计算机用户协会网络应用分会 2019 年第二十三届网络新技术与应用年会论文集. 中国计算机用户协会网络应用分会: 北京联合大学北京市信息工程重点实验室, 2019: 245-249.
- [11] He, K., Zhang, X., Ren, S. and Sun, J. (2014) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **37**, 1904-1916. <https://doi.org/10.1109/TPAMI.2015.2389824>
- [12] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2013) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 580-587. <https://doi.org/10.1109/CVPR.2014.81>
- [13] Ren, S., He, K., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [14] Liu, W., Anguelov, D., Erhan, D., et al. (2016) SSD: Single Shot MultiBox Detector. In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *Computer Vision—ECCV 2016. ECCV 2016. Lecture Notes in Computer Science*, Vol. 9905, Springer, Cham, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [15] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [16] 赵旭江. 基于卷积神经网络的遥感图像目标检测与识别[D]: [硕士学位论文]. 合肥: 中国科学技术大学, 2018.
- [17] 邓小桐,曹铁勇,方正,郑云飞. 改进 RetinaNet 的伪装人员检测方法研究[J]. 计算机工程与应用, 2021, 57(5): 190-196.

- [18] Lin, T.Y., Goyal, P., Girshick, R., *et al.* (2017) Focal Loss for Dense Object Detection. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2999-3007. <https://doi.org/10.1109/ICCV.2017.324>
- [19] 宋欢欢, 惠飞, 景首才, 郭兰英, 马峻岩. 改进的 RetinaNet 模型的车辆目标检测[J]. 计算机工程与应用, 2019, 55(13): 225-230.