

Application of Cascade R-CNN and YOLOv3 in Missile Target Recognition

Zhiying Liu¹, Chunsi Xie², Jinjun Li², Yu Sang¹

¹Midshipmen Group Five, Dalian Naval Academy, Dalian Liaoning

²Department of Missile, Dalian Naval Academy, Dalian Liaoning

Email: 18813136733@163.com

Received: Mar. 4th, 2020; accepted: Mar. 19th, 2020; published: Mar. 26th, 2020

Abstract

The development of artificial intelligence technology and computer vision technology provides a new technical support for the shipborne missile to attack all kinds of sea and land targets accurately. The automatic target recognition technology based on deep learning provides a new technical guarantee for improving the accuracy of missile target recognition. This paper introduces a variety of target recognition algorithms based on convolutional neural network and applies the YOLOv3 and Cascade R-CNN algorithm to the missile target recognition and detection experiments. The experimental results show that the two algorithms have their own advantages. The accuracy and recall rate of YOLOv3 algorithm are lower than that of Cascade R-CNN, but its detection efficiency is higher than that of Cascade R-CNN. In the process of target recognition, using deep learning algorithm is an effective way for missiles to improve the accuracy of attacking targets.

Keywords

Missile Terminal Guidance, Automatic Target Recognition, Cascade R-CNN, YOLOv3

Cascade R-CNN和YOLOv3在导弹目标识别中的应用

刘志赢¹, 谢春思², 李进军², 桑雨¹

¹海军大连舰艇学院学员五大队, 辽宁 大连

²海军大连舰艇学院导弹与舰炮系, 辽宁 大连

Email: 18813136733@163.com

收稿日期: 2020年3月4日; 录用日期: 2020年3月19日; 发布日期: 2020年3月26日

摘要

人工智能技术和计算机视觉技术的发展,为舰载导弹准确攻击各类海上、陆上目标识别提供了新的技术支持,基于深度学习的自动目标识别技术对提高导弹寻的制导精度提供了新的技术保证。介绍了多种基于卷积神经网络的目标识别算法,应用YOLOv3和Cascade R-CNN算法进行了导弹目标识别检测实验。实验结果表明,二种算法各有千秋,YOLOv3算法的准确率和召回率低于Cascade R-CNN,但其检测效率高,因此,在目标识别过程中,采用深度学习算法是导弹提高攻击目标准确性的一种有效途径。

关键词

导弹末制导, 自动目标识别, Cascade R-CNN, YOLOv3

Copyright © 2020 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

随着武器装备的发展日趋智能化,自动目标识别(Automatic target recognition, ATR)成为导弹末制导研究的重要课题[1]。自20世纪60年代提出ATR概念以来,ATR技术得到迅猛发展,相关学者对不同制导模式提出了算法理论,并伴随技术发展不断提出新算法[2]。当前,基于可见光、激光和红外图像识别等技术在导弹目标识别中已广泛应用,但识别精度有待进一步提高,迫切需要研究新的方法和技术,提高对打击目标识别的准确性和实时性。

2006年,深度学习被Geoffrey Hinton提出并迅速应用于计算机视觉领域[3],其本质是运用分层特征表示方法,对大数据样本进行分析,提取样本中的低级特征,并进一步抽象化实现高级特征表示,具有可靠性高、智能化等特点,被广泛应用于语音识别、图像识别和音频检测等领域[4]。深度学习为实现导弹打击目标自动识别提供了新的技术途径,其中性能较为优异的算法有Cascade R-CNN和YOLOv3。

2. 基于CNN网络的目标识别算法

深度学习是机器学习的一个重要分支。与传统经典机器学习算法相比,深度学习神经网络模型的中间隐层数量更多,结构更为复杂,可通过逐层训练和迭代循环,实现对复杂函数的逼近。样本数据的高级特征是通过对其低级特征进行抽象运算获取的,具备学习样本数据更为本质特征的能力[5]。当前具有代表性的网络模型有深度置信网络(deep belief network, DBN)模型、堆叠自动编码器(stacked auto-encoder, SAE)模型、基于循环神经网络(recurrent neural networks, RNN)的深度模型和基于卷积神经网络(convolutional neural network, CNN)的深度模型。

CNN网络是一种适合处理空间数据(如图像、视频)的前馈式神经网络,受视觉神经科学的启发,具备自主分层学习数据的能力[6],在计算机视觉领域广泛应用。经典CNN模型结构如图1所示,由输入层、多层卷积及池化层、光栅层、多层感知器及输出层组成,其中卷积层提取局部特征,池化层降低隐藏层维度、减少参量并使特征具有空间不变性,光栅层将多维特征线性化,多层感知器实现分类和预测。

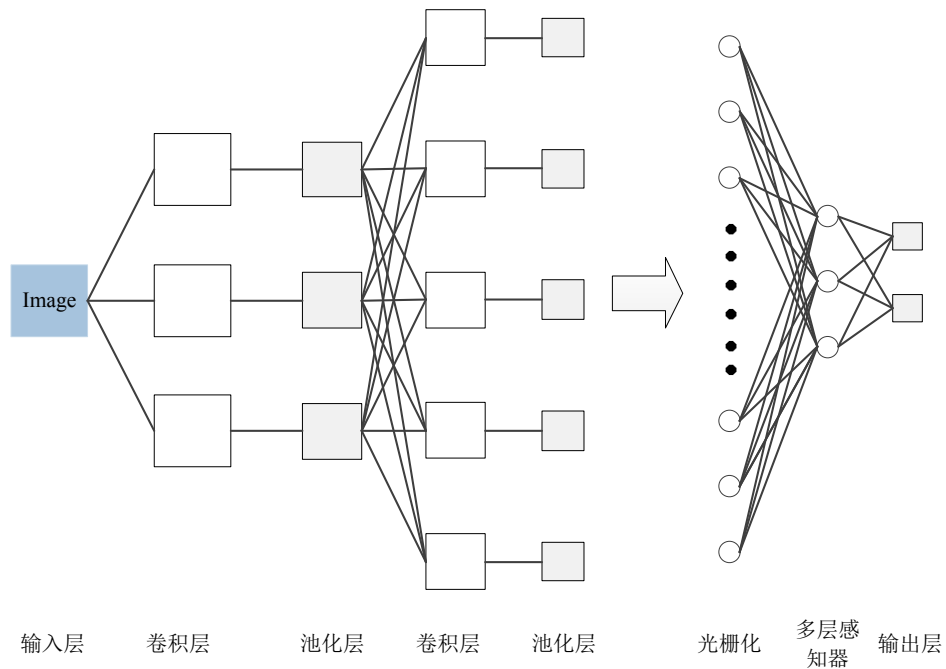


Figure 1. CNN network structure
图 1. CNN 网络结构

1989 年 Yann LeCun 首次提出 CNN 模型[7]。得益于计算机计算能力的提升，2012 年 Hinton 博士提出了深度卷积神经网络 AlexNet [8]，该网络在当年图像识别领域竞赛中的 top-1 的 error rates 为 37.5%，远超第二名的优异性能引起研究者的重视。随着 CNN 网络的不断深化，网络从大量数据中学习提取目标特征的能力不断增强，针对图像的目标识别算法研究也取得重大进展[9]。深度 CNN 目标检测算法分为两类[10]：双步检测算法，将目标识别分为分类和回归候选框两个问题去考虑，如 Fast R-CNN，Faster R-CNN，Cascade R-CNN 等，一般该类识别算法具有较高的准确性；单步检测算法，分类和候选框同时完成，如 SSD，YOLO，YOLOv2，YOLOv3 等，一般该类识别算法具有较高的检测速度。以目标识别平均准确率(mean average precision, mAP)和每秒检测帧数(frames per second, FPS)为指标，各算法在 VOC2007/2012 和 COCO 集上的表现如表 1 所示。

Table 1. Comparison of target detection performance of each algorithm
表 1. 各算法目标检测性能比较

目标识别算法	处理器	数据集	mAP	FPS
Fast R-CNN			70.0	0.5
Faster R-CNN			73.2	7
SSD500	Geforce GTX TiTan X	VOC2007 + 2012	76.8	19
YOLO			63.4	45
YOLOv2 (416)			76.8	67
SSD500			31.2	8
YOLOv2 (416)			21.6	
Faster R-CNN	Pascal TiTan X	COCO test-dev	34.9	
Cascade R-CNN			42.8	
YOLOv3 (416)			31.0	35

由表 1 可知, 在识别平均准确率方面, VOC 数据集中 SSD500 最高达到 76.8%, 但其在 COCO 集中仅为 31.2%, 与 YOLOv3 的 31.0% 相当, 低于 Cascade R-CNN 的 42.8%; 每秒检测帧数方面, VOC 数据集中 SSD500 为 19 fps, YOLOv2 (416) 最高为 67 fps, 但在 COCO 集中 YOLOv3 (416) 达到 35 fps, 远高于 SSD500 的 8 fps。基于以上分析, Cascade R-CNN 在双步检测算法中识别准确率最高, YOLOv3 在兼顾单步检测算法高速检测能力的基础上, 识别准确率有较大提升。

3. Cascade R-CNN 和 YOLOv3

3.1. Cascade R-CNN 目标识别算法

Cai Zhaowei [11] 于 2018 年在 CVPR2018 会议上提出 Cascade R-CNN 网络模型, 该模型基于 Faster R-CNN 网络的构建方法, 通过多阶段检测架构, 每阶段输入不同质量的样本数据, 并逐级提高边界框 IOU 指标阈值, 实现级联检测[12], 有效改善了 Faster R-CNN 网络检测时难以排除假阳性样本的问题[13]。

Faster R-CNN 目标识别流程如图 2 所示。首先将图像输入 CNN 网络进行特征提取, 生成特征图; 然后利用 RPN 网络提取出候选区域框, 与特征图最后一层合并生成候选区域特征; 再利用兴趣区域(ROI)池化层将候选区域特征输入到全连接层; 最后利用 SoftMax 分类器进行分类识别。

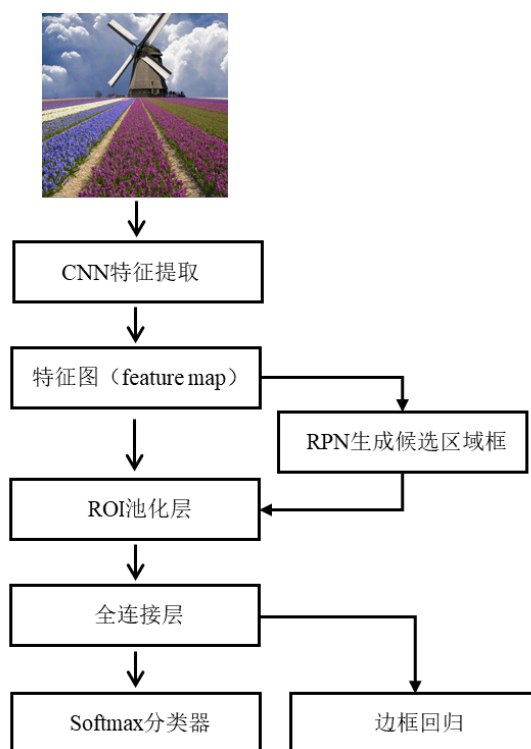


Figure 2. Fast R-CNN algorithm target recognition process

图 2. Faster R-CNN 算法目标识别流程

Cascade R-CNN 网络目标识别流程如图 3 所示。将图片输入 CNN 网络进行特征提取, 生成特征图; 然后在 Faster R-CNN (红色框内) 中进行分类和回归操作; 将回归产生的边框 1 输入 RPN 网络再次候选区域操作, 不同的是 IOU 指标阈值更高, 如此操作两次, 得到回归边框 2 和 3; 最终将所有分类结果和回归边框输出

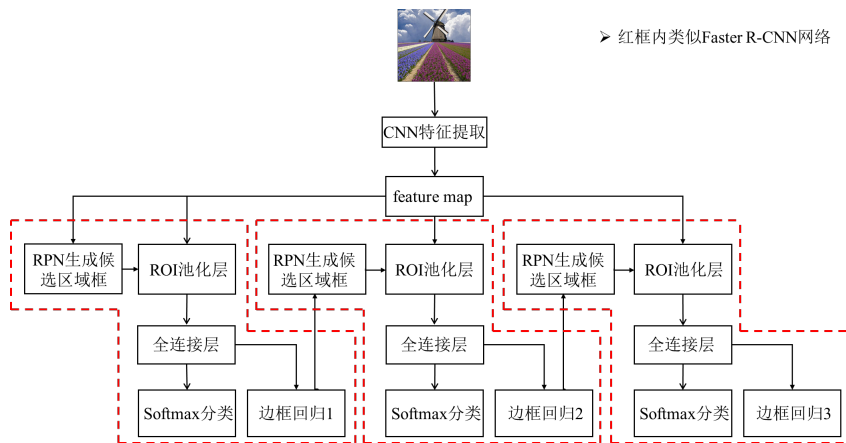


Figure 3. Cascade R-CNN algorithm target recognition process
图 3. Cascade R-CNN 算法目标识别流程

3.2. YOLOv3 目标识别算法

2016 年, Redmon 在 CVPR2016 会议上提出了一个基于深度卷积神经网络的端到端目标识别模型——YOLO (You only look once)模型[14], 并在此基础上对模型进行一系列改进, 于 2017 年提出 YOLO9000 模型, 目标识别精度大幅提升[15]。2018 年, Redmon 将残差模型和 FPN 架构引入 YOLO 提出了 YOLOv3 模型, 在保证运算效率的前提下, 进一步提高了检测精度[16]。

YOLOv3 模型的网络结构如图 4 所示, 使用 DarkNet53 模型作为目标特征提取网络。首先将输入的图像放缩到 416×416 大小, DarkNet53 网络将图像按照预先设定的特征图尺度大小(13×13 , 26×26 , 52×52)划分为 $S \times S$ 个相同尺寸的单元格, 然后将提取的浅层特征和深度特征融合获得特征金字塔, 最后再依据每个单元格每种尺度下对应的 3 个锚框(anchor box)回归预测 3 个边框(bounding box)。YOLOv3 的多尺度融合设计有效改善了 YOLO 系列算法对小尺度目标识别能力弱的缺点, 成为兼顾检测精度和速度的优异算法[17]。

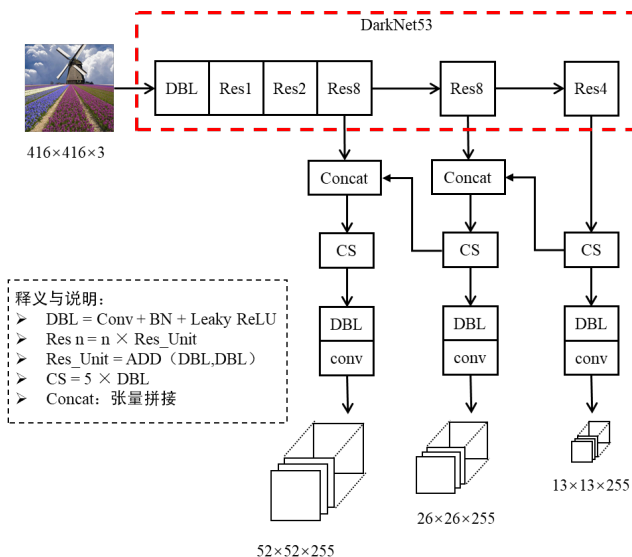


Figure 4. YOLOv3 algorithm target recognition process
图 4. YOLOv3 算法目标识别流程

4. 实验结果与分析

基于以上分析, 将两类目标识别算法中具有代表性的 Cascade R-CNN 和 YOLOv3 应用到导弹目标识别领域中, 通过目标检测实验比较不同算法的优劣性。

4.1. Cascade R-CNN 和 YOLOv3 模型训练和识别

值得注意的是, 当前算法将目标识别为不同的“种类”, 如建筑、车辆、飞机等。而在军事行动中, 每一枚导弹所要打击的目标是确定的“个体”, 而不是同种类的其他“个体”, 如建筑 A、车辆 A 等, 这就要求训练样本数据集必须源自单一待打击目标。假定某岛屿为导弹待打击目标, 本文通过搜集该岛屿的互联网图集和航拍视频, 将视频分解为图像帧, 并通过数据增广, 构建了一个包含 7500 张图像的该岛屿样本数据集用于模型的学习和训练。设定训练集和测试集的比例分别为 0.8 和 0.2, 实验平台参数为: Intel(R) Core(TM) i5-8400 CPU@2.80GHz 处理器, 8G 内存, 4G NVIDIA GeForce GTX 1050 Ti GPU。部分样本数据集如图 5 所示。



Figure 5. Sample dataset partial image
图 5. 样本数据集部分图像

为节省网络的训练时间, 加快模型的收敛速度, 利用已有权重文件进行预训练。实验中, 在 Cascade R-CNN 模型和 YOLOv3 模型训练环节, 设置循环次数为 2000, 初始学习率为 0.005, 利用 Adam 优化器, 根据返回的损失函数实时调整学习率, 循环 1500 次时调整学习率为 0.001, 避免模型陷入局部最小; 在测试环节, 分别将测试集输入到 Cascade R-CNN 模型和 YOLOv3 模型中, 当预测的目标边框与标注的边框 $\text{IOU} > 0.5$ 时认为识别正确, 否则识别错误。识别结果如图 6 和图 7 所示, 两个模型都能实现对设定岛屿的识别。



Figure 6. Result of island recognition by Cascade R-CNN
图 6. Cascade R-CNN 算法对设定岛屿识别结果

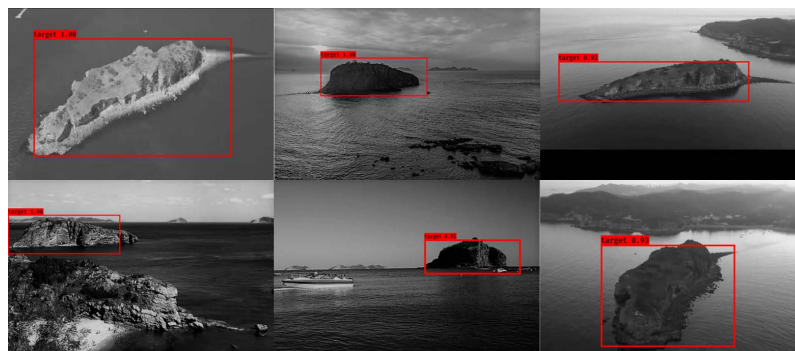


Figure 7. Result of island recognition by YOLOv3

图 7. YOLOv3 算法对设定岛屿识别结果

4.2. Cascade R-CNN 和 YOLOv3 识别性能评估

实验以准确率 P (Precision)、召回率 R (Recall)、平均精确度 AP (Average Precision)和每秒检测张数 FPS 为指标，客观评价两个模型性能。准确率表达了模型“识准”的能力，召回率表达了模型“识全”的能力，二者计算方法如式 1~2 所示，式中 TP 表示真正例，即预测结果和实际结果都为正； FP 表示假正例，即预测结果为正，实际结果为负； FN 表示假负例，即预测结果为负，实际结果为正。

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

测试集共含 1500 张 320×240 的图片，经统计包含设定岛屿目标的图片有 1454 张。Cascade R-CNN 模型正确识别该岛屿目标的个数为 1429 个，误判为岛屿目标的个数为 106 个，测试总耗时 178.5 s；YOLOv3 模型正确识别该岛屿目标的个数为 1403 个，误判为岛屿目标的个数为 311 个，测试总耗时 40.2 s；传统 HOG + SVM 目标识别算法正确识别该岛屿目标的个数为 979，误判为岛屿目标的个数为 358 个，测试总耗时超过 1 小时。两个模型与 HOG + SVM 的客观指标对比如表 2 所示，准确率 - 召回率曲线如图 8 所示。

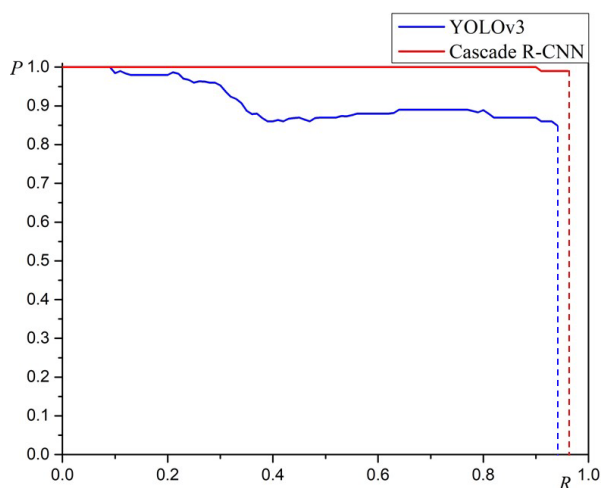


Figure 8. Performance of precision and recall

图 8. 准确率 - 召回率曲线

Table 2. Performance comparison of Cascade R-CNN and YOLOv3 in missile target recognition
表 2. Cascade R-CNN 和 YOLOv3 算法在导弹目标识别中的性能比较

目标识别算法	P	R	AP	FPS
HOG + SVM	73.22%	67.33%	53.84%	<0.42
Cascade R-CNN	93.1%	98.3%	95.6%	8.4
YOLOv3	81.9%	96.5%	88.7%	37.3

分析表 2 和图 8 可知, 针对设定的导弹打击某岛屿目标, 与传统机器学习目标识别算法相比, 基于 Cascade R-CNN 和 YOLOv3 算法的目标识别模型准确性和实时性均有极大提高。其中, Cascade R-CNN 算法的目标识别模型准确率、召回率和 AP 值分别比基于 YOLOv3 算法的目标识别模型高 11.2%、1.8% 和 6.9%; 检测速度方面, YOLOv3 模型每秒检测帧数达 37.3 fps, 是 Cascade R-CNN 模型(8.4 fps)的 4.4 倍。Cascade R-CNN 算法的准确性更高, YOLOv3 算法实时性更好。

5. 总结

针对导弹目标识别技术的现实需求, 介绍了近年来深度学习网络模型的发展状况和基于深度卷积神经网络的目标识别算法, 将具有代表性的 Cascade R-CNN 和 YOLOv3 算法应用到导弹目标识别领域并进行了比较, 得出如下结论:

1) 基于深度学习算法的目标识别模型性能远远优于基于传统机器学习的识别算法;

2) 基于 Cascade R-CNN 算法的目标识别模型的准确率和召回率均比基于 YOLOv3 算法的目标识别模型高;

3) YOLOv3 模型的 FPS 是 Cascade R-CNN 模型的 4.4 倍, 检测效率远高于 Cascade R-CNN 模型。

后续工作将结合 Cascade R-CNN 算法的准确性和 YOLOv3 算法的快速性, 兼顾检测速度和精度, 研究新的目标识别算法, 使之更适用于导弹目标识别领域。

参考文献

- [1] 赵晓震. 导弹自动目标识别关键技术研究[J]. 飞航导弹, 2014(2): 51-56.
- [2] 江言林, 高立志, 刘铁军. 反舰导弹目标识别选择技术的发展应用[J]. 飞航导弹, 2009(1): 53-55.
- [3] Hinton, G.E., Osindero, S. and The, Y.W. (2006) A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, **18**, 1527-1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
- [4] 马啸, 邵利民, 金鑫, 徐冠雷. 深度学习技术及其在舰船目标识别领域的应用[J]. 计算机技术与发展, 2019, 29(12): 141-147.
- [5] 田启川, 王满丽. 深度学习算法研究进展[J]. 计算机工程与应用, 2019, 55(22): 25-33.
- [6] 张荣, 李伟平, 莫同. 深度学习研究综述[J]. 信息与控制, 2018, 47(4): 385-397 + 410.
- [7] Lecun, Y., Boser, B.E., Denker, J.S., et al. (1989) Back Propagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, **1**, 541-551. <https://doi.org/10.1162/neco.1989.1.4.541>
- [8] Krizhevsky, A., Sutskever, I. and Hinton, G. (2012) ImageNet Classification with Deep Convolutional Neural Networks.
- [9] 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.
- [10] 曹昭睿, 白帆, 刘凤丽, 郝永平. 基于轻量化神经网络的目标识别跟踪算法研究[J/OL]. 弹箭与制导学报, 1-5. <http://kns.cnki.net/kcms/detail/61.1234.tj.20191211.1730.006.html>, 2020-03-24.
- [11] Cai, Z.W. and Nuno, V. (2019) Cascade R-CNN: High Quality Object Detection and Instance Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/TPAMI.2019.2956516>
- [12] 周刚. 基于 Cascade R-CNN 的车辆目标检测方法[J]. 内燃机与配件, 2019(20): 146-148.

- [13] 徐昕军, 江波, 黄启迪. 基于 Cascade R-CNN 的铁路路基翻浆冒泥病害智能识别方法[J]. 铁道建筑, 2019, 59(12): 99-104.
- [14] Redmon, J., Divvala, S., Girshick, R., *et al.* (2016) You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of Computer Vision and Pattern Recognition*, Boston, MA, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [15] Redmon, J. and Farhadi, A. (2016) YOLO9000: Better, Faster, Stronger. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 6517-6525. <https://doi.org/10.1109/CVPR.2017.690>
- [16] Redmon, J. and Farhadi, A. (2018) Yolov3: An Incremental Improvement.
- [17] 陈正斌, 叶东毅, 朱彩霞, 廖建坤. 基于改进 YOLOv3 的目标识别方法[J]. 计算机系统应用, 2020, 29(1): 49-58.