

基于伪标签的半监督安全帽佩戴实时检测方法

杜先锋¹, 王佳英^{1,2}

¹沈阳建筑大学计算机科学与工程学院, 辽宁 沈阳

²沈阳工业大学软件学院, 辽宁 沈阳

收稿日期: 2022年12月12日; 录用日期: 2023年1月12日; 发布日期: 2023年1月20日

摘要

在金属制造、桥梁隧道工程、建筑行业等过程中, 佩戴安全帽可以极大地保护生命安全。目标检测方法可用于检测头盔是否佩戴。但目前的安全帽佩戴检测方法多集中于监督学习, 依赖于大量精确标记的数据。但在现实中, 标记数据的成本非常高, 训练数据的获取不足可能成为性能提升的瓶颈。与有标签的数据相比, 无标签的数据更丰富、更便宜、更容易获得。基于这一问题, 将伪标签技术引入到传统安全帽检测方法中, 提出了一种半监督安全帽检测方法。它在训练模型时同时使用有标签的数据和无标签的数据, 只需要少量的有标签的数据, 而使用大量的无标签数据来辅助模型的训练。在自制头盔数据集上的实验结果表明, 该方法能在有限的标记数据下取得良好的性能, 准确率达到92.7%, 平均准确率提高3.7%。在标记数据不足的情况下, 满足头盔检测的要求。

关键词

安全头盔检测, 半监督, 伪标签

A Real Time Detection Method for Semi-Supervised Safety Helmet Wearing Based on Pseudo Label

Xianfeng Du¹, Jiaying Wang^{1,2}

¹School of Computer Science and Engineering, Shenyang Jianzhu University, Shenyang Liaoning

²School of Software, Shenyang University of Technology, Shenyang Liaoning

Received: Dec. 12th, 2022; accepted: Jan. 12th, 2023; published: Jan. 20th, 2023

Abstract

In the process of metal manufacturing, bridge and tunnel engineering, and construction industry,

文章引用: 杜先锋, 王佳英. 基于伪标签的半监督安全帽佩戴实时检测方法[J]. 数据挖掘, 2023, 13(1): 67-74.

DOI: 10.12677/hjdm.2023.131007

wearing a safety helmet can greatly protect the safety of life. The target detection method can be used to detect whether a helmet is worn or not. The current safety helmet wearing detection methods mostly focus on supervised learning, which relies on a large number of accurately labeled data. However, in reality, the marked data is very expensive, and the insufficient acquisition of training data may become a bottleneck for performance improvement. Compared with labeled data, unlabeled data are more abundant, cheaper and easier to obtain. Based on this problem, this paper introduces the pseudo-label technology into the traditional safety helmet detection method, and proposes a semi-supervised safety helmet detection method. It utilizes both labeled and unlabeled data when training the model, and it requires only a small amount of labeled data, while assisting the training of the model with a large amount of unlabeled data. The experimental results on the self-made helmet data set show that this method can achieve good performance under limited labeled data, with an accuracy rate of 92.7% and an average accuracy increase of 3.7%. It meets the requirements for helmet detection in case of insufficient marking data.

Keywords

Safety Helmet Detection, Semi-Supervised, Pseudo-Label

Copyright © 2023 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

在建筑施工、桥梁施工、金属制造等高危行业的工业生产施工现场施工中, 安全操作是保证工人生命安全的重要措施, 正确佩戴安全帽是安全操作的有力保障。但在真实的场景中, 很多员工缺乏安全意识, 不遵守佩戴安全帽的规则, 在发生事故时造成了严重的损失。在施工现场, 经常使用人工监督来检测员工是否正确佩戴安全帽。但由于现场活动面积大, 情况复杂, 效果并不理想。因此, 开发相应的自动监控方法代替人工监控是合适的解决方案。

传统的安全帽佩戴检测基于传感器等专门的元件, 如文献[1]是基于移动射频识别(RFID)的检测方法。这种方法有一定的效果, 然而它们检测范围有限, 并且检测效果较差, 大多数只能检测头盔是否靠近工人, 但不能确认头盔是否被戴着。随着计算机显卡计算能力和存储等硬件成本不断下降, 基于神经网络的目标检测在近 10 年来取得了重大的成果。目前, 最新的目标检测框架和一些改进框架(如[2][3])已经应用到头盔检测中。[4]使用单阶段 yolov5 检测网络, 通过现场拍摄和网上采集的数据集训练得到一个针对安全帽检测的模型。[5][6]对 yolov5 方法进行改进, 前者改变其特征提取网络, 以此来加强对特征的提取; 后者优化原方法中边界框损失上置信度损失, 提高对小目标特征的学习效果。以此来提升检测精度。[7]针对安全帽识别的特点专门为其设计了一种检测多尺度检测网络, 优化了安全帽图片特征的提取和检测。[8]对 FasterRCNN 方法进行改进, 改变其特征提取网络, 针对安全帽在图片中占幅较小的特点, 优化选择的尺寸, 提高模型的检测精度。

以上基于深度学习的方法取得了一定的效果。但是, 它们依赖于大量的标记数据, 当标记数据稀缺或数据分布不同时, 效果不佳。然而, 在现实中, 我们没有这么多的标记数据, 数据稀缺已经成为性能改进的瓶颈。

基于这个问题本文提出一种基于伪标签的半监督检测方法。主要贡献如下:

1) 该方法引入了伪标记机制来利用大量未标注的信息, 让其参与到模型的训练中来。

2) 引入 `imgmix` 增强方法; 在训练模型时, 使用大量的无标签数据参与模型训练, 提高模型对分布在不同领域的数据的适应性。该方法避免了获取注释数据的困难, 以更经济的方式获得了性能优异的模型。

3) 收集数据, 制作安全帽数据集, 对本文提出的方法进行实验, 验证其性能。

2. 相关工作

2.1. 目标检测算法

近年来, 随着深度学习的发展和计算机硬件的进步, 利用目标检测进行这类检测成为一种新的选择。物体检测的任务是在图像中找到所有感兴趣的物体, 确定它们的类别和位置。最新的目标检测研究方法主要分为两大类; 它们是一级方法(例如, YOLO [9], SSD [10], Retinanet [11], RefineDet [12])和二级方法(例如, fast/faster R-CNN [13] [14], FPN [15], MaskR-CNN [16])。两阶段算法首先提出一个建议框, 第一步通过网络返回目标框前景的大致位置、大小和概率, 第二步通过另一个网络返回目标框的位置、大小和类别; 一级网络的核心是通过网络对输入的图像直接返回目标的大小、位置和类别。它们依赖于大量的标记数据, 但现实生活中制作数据集成本非常昂贵, 这个因素成为目标检测框架性能进一步发展的桎梏。

2.2. 伪标签

基于伪标签的学习方法是半监督学习的一种。半监督学习方法有两种: 一致性方法[17]和伪标签方法[18]。本文中使用的属于伪标签的范畴。

伪标签技术最早是 Lee 等人提出, 他们考虑到现实中有标注的数据集数据量难以满足需求且标注成本昂贵但无标注的数据极其丰富, 且获取成本低廉的情况, 开创性的提出同时利用有详尽标注的数据和没有标注无标签的数据来进行模型的训练。对于没有进行人工标记的数据, 利用在有标注数据集上得到的模型对其生成伪标签(Pseudo-Labels)并且把生成的伪标签当成真实标签来使用, 生成伪标签时它只选取那些具有预测概率得分最高的类。这实际上相当于熵正则化。经过不断发展和完善, 伪标签技术现今被广泛应用在机器学习和深度学习的各个领域。

3. Pseudo-Label-Based 检测方法

本文方法选用 `yolov5` 模型做为本方法的基线方法, 首先使用 `YOLOv5` 模型在有详细标注的源域数据集 $D_s = \{(X_i, Y_i)\}$ 上进行训练, X_i 代表数据集的图像, Y_i 代表与其相对应的标签。 Y_i 标签中包括图像中目标物体的类别信息, 还有目标物体标注框四个角的坐标。在训练中我们要将没有标注的目标域数据集 $D_T = \{(X_i)\}$ 中的数据利用起来, 使用伪标签技术让其也参加模型的训练, 以助于模型学习未标注样本的域特征, 提升模型的检测精度。

3.1. 整体方法

安全帽检测既要求精度又要求速度, 能适应施工现场恶劣的环境。为了满足这些要求, 我们必须首先拥有大量精确标记的数据集。但是手工标记大数据集的成本是昂贵的, 甚至比施工期间使用工人现场监督的成本还要高。相反, 网络中有大量未被利用的原始(未标记的)数据。针对这种情况, 我们提出了一种半监督安全帽检测方法。它使用大量的未标记数据帮助训练模型, 在生成伪标签时只考虑精度, 使用精度较高的模型生成伪标签, 并对伪标签进行细化; 当实际部署应用程序时, 将使用一个更快的模型。

方法流程如图 1 所示:

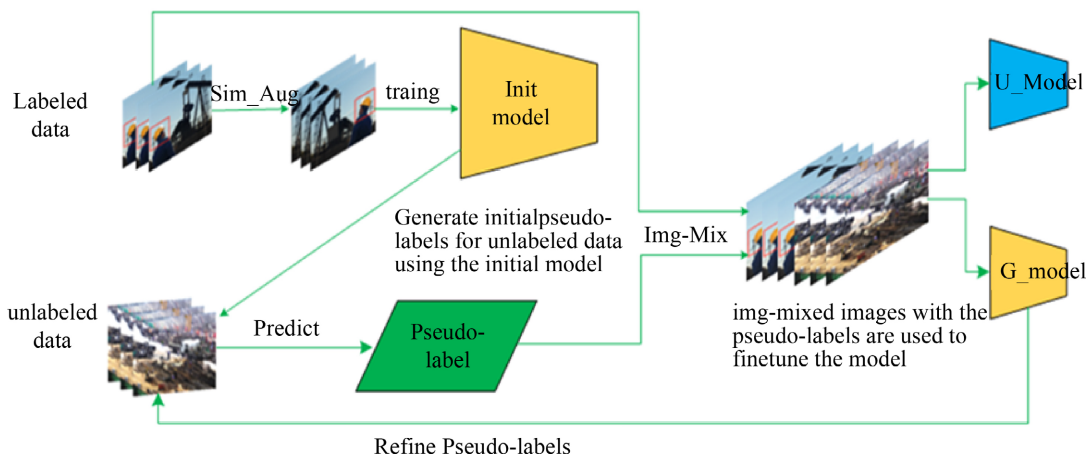


Figure 1. Overview flow

图 1. 总体流程

- 步骤 1. 使用 YOLOv5 模型(Init model)当作基准模型在精确标记的源域数据集上进行训练, 得到在源域数据集上的权重参数。在这一步中对输入的图片数据进行简单的增强, 如旋转、裁剪、高斯噪声等。然后使用在源域数据集上训练得到 YOLOv5 模型在没有标记的目标域数据上生成初始的伪标签, 包括目标物体类别和边界框四个角的坐标信息(X_1, Y_1, X_2, Y_2)等。
- 步骤 2. 将带有伪标签的目标域数据使用 ImgMix 方法(下面会对这个方法进行介绍)与带有真实标签的源域数据进行混合和增强, 得到一个混合后的数据集。在混合后的数据集上重新训练 yolov5 模型, 此时用到混合数据集含有目标域的特征信息, 所以在此数据集上训练的模型能够学习到目标域数据的特征。
- 步骤 3. 使用混合域中训练的 yolov5 模型(G_model)在未标记的目标数据上重新生成伪标签。
- 步骤 4. 循环执行步骤 2 和步骤 3, 直到模型收敛(U_model)。

3.2. ImgMix

在我们的方法中, 我们使用一种名为 ImgMix 的方法进行数据增强, 如图 2 所示, 它从标记的和未标记的数据中采样, 将采样的图像及其标签(伪标签)混合到新图像中。该方法与其他增强方法的区别主要体现在两个方面:

- 它要求裁剪区域必须来自目标域, 并将裁剪部分与源域图像相结合。
- 源域图像的显著性区域在图像混合过程中不能被破坏。

具体步骤如下:

步骤 1: 它从标记数据和未标记数据中随机抽样。随机取未标记数据样本中的所有预测边界框中的一个或多个, 粘贴到已采样标记数据的样本上, 替换原来大小相同的部分, 形成一组新的图像。

步骤 2: 它在一幅图像上混合了精确的注释和伪标签, 使模型能够更好地学习无标签数据的特征。

4. 实验分析

4.1. 数据集采集与处理

本文的数据集主要来源于网络抓取、公共数据集清理和视频帧提取。该数据集包含 14,205 张图像,

分为两部分,源数据和目标数据,分别为 5000 和 9205 张。对源数据进行精确的手工标记(标记为 labelImg,戴帽者和不戴头盔者标记为人员),并对大小和格式进行调整;目标数据没有标记,只有简单的大小和格式调整(调整为 JPG 格式)。这次收集的样本数据集如图 3 所示:

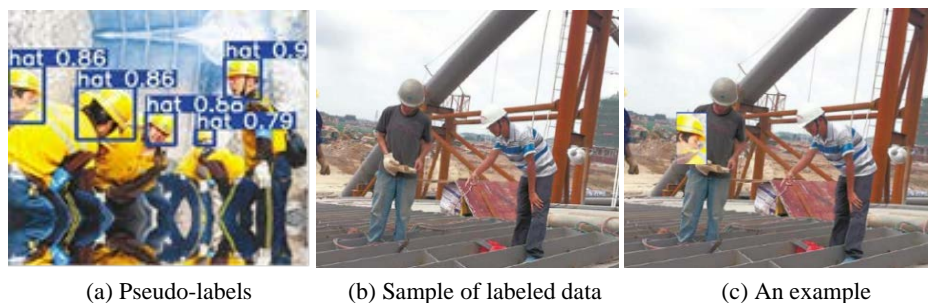


Figure 2. ImgMix Example
图 2. ImgMix 例子

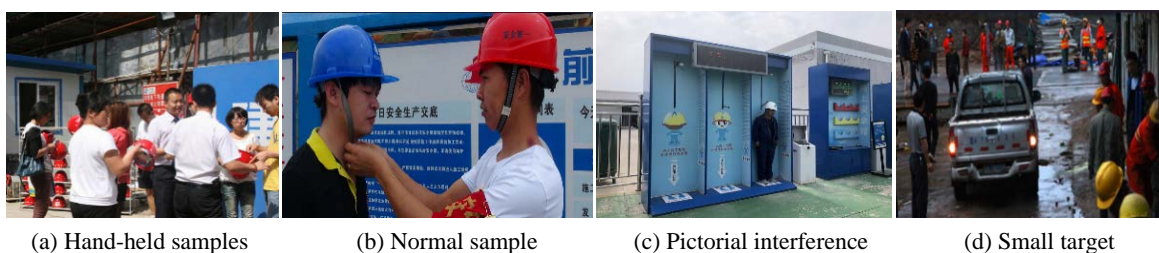


Figure 3. Sample data set example
图 3. 样本数据集示例

4.2. 评价指标

召回率、精度、平均精度(AP)、平均平均精度(mAP)、帧每秒(FPS)是目标检测中常用的评价指标,其计算公式如下:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{AP} = \sum_{i=1}^{n-1} (r_{i+1} - r_i) P(r_{i+1}) \quad (3)$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (4)$$

以人为例解释涵义: TP (true Positive)指被预测为人的; FP(假阳性)指被预测为人的帽子; FN (假阴性)指的是被预测是假阴性的人。AP 是指城市被分别由 x 和 y 的召回轴和精度轴组成的曲线所包围的区域。mAP 是 AP 的重平均,是所有类别 AP 的平均值。

4.3. 实验环境配置

为了验证本文半监督头盔检测方法的准确性和有效性,在参数如表 1 所示的环境平台下进行了实验。

Table 1. Experimental environment and configuration**表 1.** 实验环境与配置

实验环境	配置
System	windows 11
CPU	Intel 12700H
内存	16G
GPU	RTX3060
Python	3.8
Editor	Pycharm
框架	Pytorch (1.11.0)
cuda	11.6

4.4. 实验结果与分析

为了验证本文提出的头盔检测方法是否具有良好的性能, 使用相同的数据集和配置对几种流行的检测框架进行了实验, 通过搭建和自身烧蚀实验对 `ImgMix` 的效果进行了验证。其中有单阶段方法 `SSD` 和 `yolov5`, 两阶段方法 `Faster R-CNN`。采用 `AP` 和 `mAP` 指标进价。

从表 2 的结果可以看出, 本文提出的半监督安全帽检测方法可以有效提高建筑工人是否戴安全帽的检测精度, 同时兼顾了检测速度。在检测准确率方面, 与原有的 `Faster R-CNN`、`SSD`、`YOLOv5` 等方法相比, 该方法对佩戴安全帽的施工人员的平均准确率 94.1%; 该方法对无安全帽施工人员的检测平均正确率为 91.3%; `mAP` 达到 92.7%, 优于原方法。在检测速度方面, 我们的 `U_model` 选择了 `YOLOv5` 模型, 所以速度和它是一样的, 可以满足实时检测的要求。

Table 2. Experimental result**表 2.** 实验结果

Method	AP50/%		mAP/%	Inferring time/ms
	hat	person		
Faster R-CNN	81.8	41.2	61.5	289
SSD	78.8	68.2	73.5	122
YOLOv5	89.3	88.6	89	36
Ours	94.1	91.3	92.7	37

为了验证各改进模块的优化效果, 设置了以下烧蚀对比实验。结果如表 3 和图 4 所示, 其中图 4 是一个 `IOU` 为 0.5 的 `PR` 曲线, 其中灰色线表示底层的 `yolov5` 方法, 橙色线表示使用伪标签方法, 蓝色线表示使用伪标签方法和 `ImgMix` 方法。

结合表 3 和图 4, 我们可以看到每个部分的作用(图 4 中 `T` 代表教师指导的伪标签, `IM` 代表 `ImgMix` 图像增强方法)。仅使用教师指导伪标签(`T-PL`, 以下由 `T-PL` 代替), 和伪标签与 `ImgMix` 图像增强方法都使用得到的 `mAP` 与原始方法相比分别增加了 2.5 和 3.7 个百分点。从表 3 可以看出, 使用 `T-PL` 方法后, 模型对戴头盔人员的检测准确率提高了 3.3 个百分点, 对不戴头盔人员的检测准确率提高了 1.8 个百分点。之后使用 `T-PL` 和 `ImgMix` 方法, 对戴头盔人员的模型检测准确率比原方法提高了 4.8 个百分点, 对不戴

头盔人员的模型检测准确率提高了 2.7 个百分点。这可以说明, 将大量的未标记数据应用到模型训练中, 提高硬帽检测方法的准确性是可行的, 极大地缓解了人工标记数据成本高、数量不足的压力。

Table 3. Results of ablation experiment

表 3. 消融实验结果

Method	T-PL	ImgMix	AP50/%		mAP/%
			hat	person	
YOLOv5	×	×	89.3	88.6	89
Ours	√	×	92.6	90.4	91.5
Ours	√	√	94.1	91.3	92.7

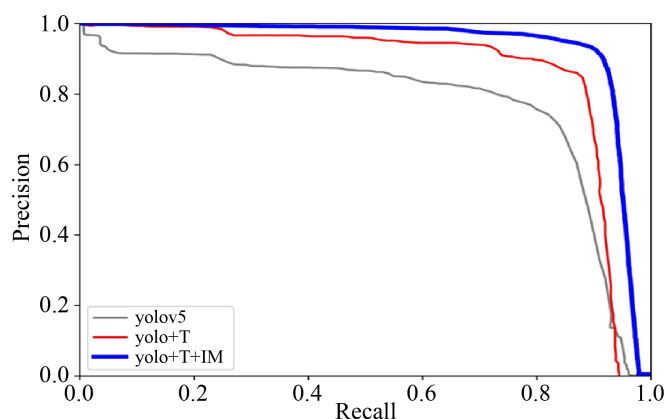


Figure 4. PR curve

图 4. PR 曲线

5. 结论与展望

本文提出了一种用于半监督学习的安全帽检测方法。首先, 利用人工精密标签源数据训练初始模型, 利用该模型获得初始伪标签; 其次, 利用 `ImgMix` 方法将带标签的源数据和带伪标签的无标签数据进行混合, 充分融合有标签数据和无标签数据的特征, 对模型进行再训练, 调整伪标签; 最后, 利用最终的伪标签训练出一个检测速度非常快的框架。从实验结果可以看出, 本文提出的方法能够获得较好的检测精度和速度, 基本能够满足工作现场头盔佩戴的实时检测。但是, 方法操作过程比较复杂, 不是一个端到端框架, 因此下一步的研究重点是改进过程, 设计一个端到端框架, 使操作更加友好。

参考文献

- [1] Kelm, A., Meins-Becker, A., *et al.* (2013) Mobile Passive Radio Frequency Identification (RFID) Portal for Automated and Rapid Control of Personal Protective Equipment (PPE) on Construction Sites. *Automation in Construction*, **36**, 38-52. <https://doi.org/10.1016/j.autcon.2013.08.009>
- [2] Huang, L., Fu, Q., He, M., Jiang, D. and Hao, Z. (2021) Detection Algorithm of Safety Helmet Wearing Based on Deep Learning. *Concurrency Computation Practice and Experience*, **33**, e6234. <https://doi.org/10.1002/cpe.6234>
- [3] Long, X., Cui, W. and Zheng, Z. (2019) Safety Helmet Wearing Detection Based on Deep Learning. 2019 *IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chengdu, 15-17 March 2019, 2495-2499. <https://doi.org/10.1109/ITNEC.2019.8729039>
- [4] 史德伟, 郭秀娟. 基于 YOLOv5 的安全帽检测研究[J]. 吉林建筑大学学报, 2022, 39(5): 85-88.

- [5] 杜晓刚, 王玉琪, 晏润冰, 古东鑫, 张学军, 雷涛. 基于 YOLO-ST 的安全帽佩戴精确检测算法[J]. 陕西科技大学学报, 2022, 40(6): 177-183+191. <https://doi.org/10.19481/j.cnki.issn2096-398x.2022.06.004>
- [6] 吕宗喆, 徐慧, 杨骁, 王勇, 王唯鉴. 面向小目标的 YOLOv5 安全帽检测算法[J/OL]. 计算机应用, 1-9. <http://kns.cnki.net/kcms/detail/51.1307.TP.20220929.1425.005.html>, 2022-12-13.
- [7] 钱大龙, 韦古强, 叶良浩. 一种基于特征融合的安全帽佩戴识别方法[J]. 自动化技术与应用, 2022, 41(11): 69-72. [https://doi.org/10.20033/j.1003-7241.\(2022\)11-0069-04](https://doi.org/10.20033/j.1003-7241.(2022)11-0069-04)
- [8] 朱玉华, 杜金月, 刘洋, 颜永鹏. 基于改进 Faster R-CNN 的小目标安全帽检测算法研究[J]. 电子制作, 2022, 30(19): 64-66+83. <https://doi.org/10.16589/j.cnki.cn11-3571/tn.2022.19.004>
- [9] Redmon, J., Divvala, S.K., Girshick, R.B. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [10] Liu, W., Anguelov, D., Erhan, D., et al. (2016) SSD: Single Shot Multibox Detector. *Computer Vision-ECCV 2016 14th European Conference*, Amsterdam, 11-14 October 2016, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [11] Lin, T., Goyal, P., Girshick, R.B., et al. (2017) Focal Loss for Dense Object Detection. *IEEE International Conference on Computer Vision, ICCV 2017*, Venice, 22-29 October 2017, 2999-3007. <https://doi.org/10.1109/ICCV.2017.324>
- [12] Zhang, S., Wen, L., Bian, X., et al. (2017) Single-Shot Refinement Neural Network for Object Detection. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 4203-4212. <https://doi.org/10.1109/CVPR.2018.00442>
- [13] Girshick, R. (2015) Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, 7-13 December 2015, 1440-1448. <https://doi.org/10.1109/ICCV.2015.169>
- [14] Ren, S., He, K., Girshick, R.B. and Sun, J. (2015) Faster R-CNN: Towards Real Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015*, Montreal, 7-12 December 2015, 91-99.
- [15] Lin T., Dollár, P., Girshick, R.B., et al. (2017) Feature Pyramid Networks for Object Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Honolulu, 21-26 July 2017, 936-944. <https://doi.org/10.1109/CVPR.2017.106>
- [16] He, K., Gkioxari, G., Dollár, P. and Girshick, R.B. (2017) Mask R-CNN. *IEEE International Conference on Computer Vision, ICCV 2017*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/ICCV.2017.322>
- [17] Jeong, J., Lee, S., Kim, J. and Kwak, N. (2019) Consistency-Based Semi-Supervised Learning for Object Detection. In: *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Curran Associates Inc., Red Hook, Article 965, 10759-10768.
- [18] Lee, D.H. (2013) Pseudo-Label: The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks. *Workshop on Challenges in Representation Learning*, Vol. 3, 896.