

基于改进Faster-RCNN的矿工检测

林哲^{1*}, 陈平华¹, 申建芳¹, 王越超²

¹广东工业大学, 计算机学院, 广东 广州

²广州城市职业学院, 机电工程学院, 广东 广州

收稿日期: 2022年3月12日; 录用日期: 2022年4月13日; 发布日期: 2022年4月20日

摘要

针对Faster RCNN仅使用单一全连接层用于目标分类和边界框回归进而导致检测准确率不高的问题, 提出了基于改进Faster RCNN的矿工检测框架, 即将目标分类问题和边界框回归问题分开处理。为了实现矿工出入计数功能, 在改进Faster RCNN框架的识别后, 设计了基于三线的矿工出入检测算法。在真实的井下数据集上实验表明, 改进Faster RCNN检测框架有更高的查准率和查全率。

关键词

Faster RCNN, 人员识别, 矿工检测, 分类和回归

Miner Detection Based on Improved Faster-RCNN

Zhe Lin^{1*}, Pinghua Chen¹, Jianfang Shen¹, Yuechao Wang²

¹School of Computer, Guangdong University of Technology, Guangzhou Guangdong

²Guangzhou City Vocational College, Guangzhou Guangdong

Received: Mar. 12th, 2022; accepted: Apr. 13th, 2022; published: Apr. 20th, 2022

Abstract

Aiming at the problem that the detection accuracy is not high when the single full connection layer is only used for target classification and bounding box regression by Faster RCNN, a miner detection framework based on improved Faster RCNN is proposed, which deals with target classification and bounding box regression separately. In order to realize the function of miner in and out counting, after improving the recognition of Faster RCNN framework, a miner in and out detection

*第一作者。

algorithm based on three lines was designed. Experiments on real downhole data sets show that the improved Faster RCNN detection framework has higher accuracy and recall.

Keywords

Faster RCNN, Pedestrian Detection, Miner Target Detection, Classification and Localization

Copyright © 2022 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

井下矿工作业是一个高危工作。井下作业矿工精确检测与统计不仅可实现对作业矿工的考勤，而且可避免矿工滞留甚至被困等问题，对保障井下矿工的安全具有重要意义。

提升机罐笼是地面与井下之间重要的运输设备，是矿井工人和施工材料进出矿井的唯一的运输工具。传统的矿工人数统计在罐笼出入口进行，采取人工清点的方法，这种方法不仅费时费力，而且容易出错。随着人工智能与计算机图像处理技术的发展，基于图像处理技术的矿工自动检测与统计已经成为了业界新的发展趋势。

矿井作业矿工必须佩戴安全帽。在矿工进出提升机罐笼时，罐笼出入口 45°斜上方安装了摄像头，这时每个矿工的安全帽没有出现任何遮挡，而且由于安全帽颜色鲜艳，特征明显，很适合作为矿工目标检测与统计的对象，如图 1 所示。



Figure 1. Target image of cage underground outlet

图 1. 罐笼井下出口图像

传统的矿工安全帽检测采用机器学习[1]的方法，需要人为地设计特征，例如 Haar-like 特征[2]、HOG 特征[3]等。近年来，随着深度学习的快速发展，传统的机器学习方法渐渐被基于神经网络的深度学习方法所替代[4]。吴冬梅等提出了基于改进 Faster RCNN 的安全帽检测算法[5]。曹燕等提出了一种基于等级 SSD 的建筑工人安全帽佩戴检测方法[6]。肖体刚等提出改进 YOLOv3 的安全帽佩戴检测方法[7]。但现阶段大多数研究都统一用全连接层进行分类和预测，没有考虑到分类和预测边界框任务之间的区别。最新的研究显示：全连接层适合分类任务，而卷积层更适合预测边界框任务[8]。

针对安全帽目标检测中没有考虑分类和预测边界框任务之间的区别而影响分类准确率问题, 本文基于 Faster RCNN 算法框架, 对 ROIpooling 层[9]的输出进行了分别处理, 分类任务用全连接层进行处理, 边界框回归任务用卷积层进行处理, 取得了理想的效果。

2. 传统的 Faster RCNN

传统的 Faster RCNN 是一个 two-stage (两阶段)网络, 网络结构如图 2 所示。

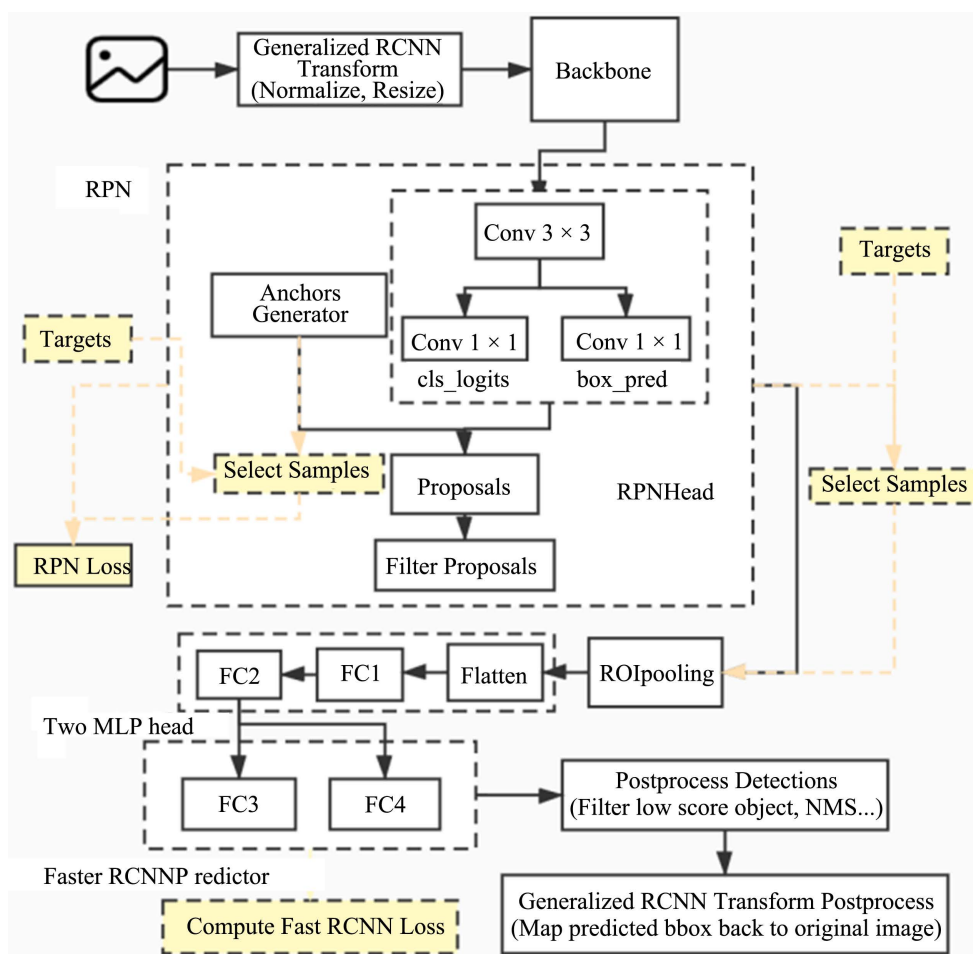


Figure 2. Traditional fast RCNN network structure

图 2. 传统的 Faster RCNN 网络结构

原始图像经过直方图均衡化预处理和 Resnet 网络[10]后首先进入 RPN (region proposal network, 区域建议网络)模块处理产生候选区域, 即预先产生系列安全帽目标候选框; RPN [11]输出经 ROIpooling 后进入目标检测模块处理, 在筛选过后的候选区域中进行目标预测, 包括展平(将二维特征图展开成一维向量)、两个全连接、目标分类以及边界框回归等子模块。

3. ResNet 网络

图 2 中的主干网络我们采用 ResNet 网络。ResNet 网络在网络中增加了直连通道, 在此之前的网络只是对输入做一个非线性变换, 而 ResNet 网络还考虑前几层的输入。考虑到原始输入的影响, ResNet 网络基础功能块 Bottleneck 如图 3 所示。

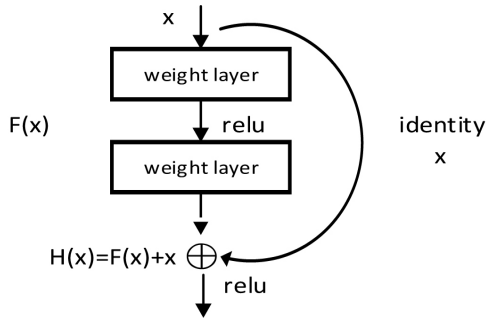


Figure 3. The realization of the ResNet
图 3. ResNet 的 Bottleneck 实现

下一层的神经网络学习上一个网络输出的残差，所以 ResNet 又叫做残差网络.提出残差学习的思想.传统的卷积网络或者全连接网络在前向传递的时候存在信息丢失，损耗等问题，同时还有梯度消失或者梯度爆炸问题，所以在此之前的网络层数都不大。ResNet 的出现使网络层数可以超过百层，通过直接将输入信息绕道传到输出，保护信息的完整性，整个网络只需要学习输入、输出差别的那一部分，简化学习目标和难度。Resnet 网络结构如图 4 所示。

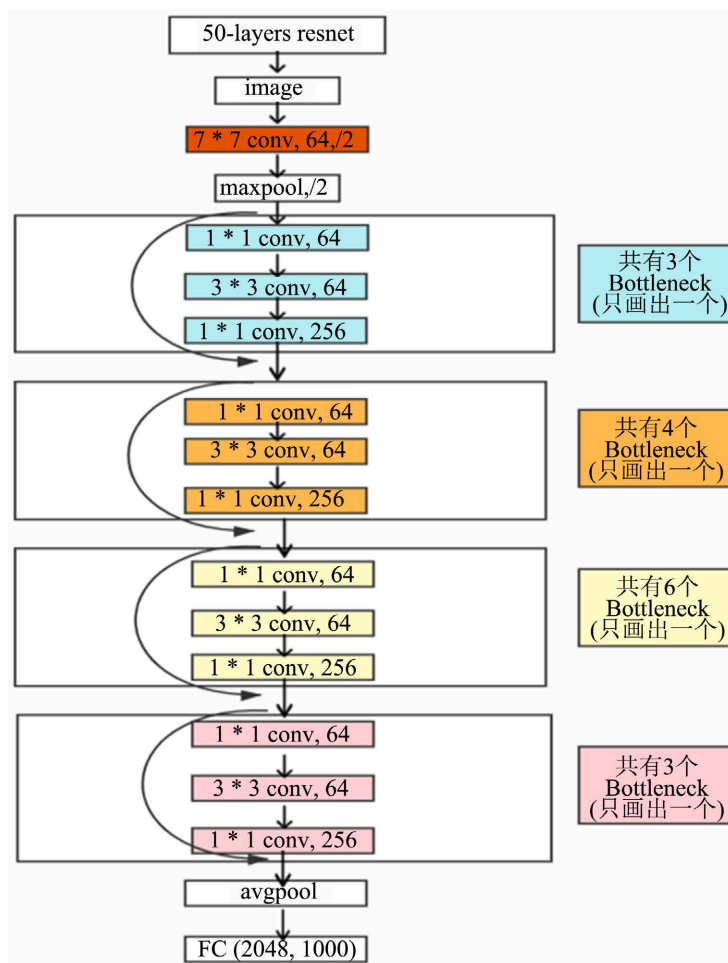


Figure 4. ResNet network architecture
图 4. ResNet 网络结构

4. 改进 Faster RCNN 的安全帽检测

传统的 Faster RCNN 的 second stage—目标检测模块在分类和边界框回归上仅使用了全连接层。对于分类和边界框回归这两个任务，仅仅使用全连接层或者卷积层往往得不到理想的效果。全连接层和卷积层在目标分类和定位两个检测任务是互补。全连接层其分类得分更适合于分类任务，因为它分类得分与建议框和真实框的交并比(IoU) [12]更相关。同时，卷积层提供了更精确的边界框回归。这是因为全连接层是空间敏感的，对于输入的不同部分有不同的参数，而卷积层对所有部分则共享卷积核。因此，全连接层能够更好地区分一个完整的对象和一个对象的一部分，而卷积层能够更稳健地回归整个对象(边界框回归)。基于以上发现，从而提出了对 ROIpooling 层输出分开处理的改进，包括一个用于分类的全连接层和一个用于边界框回归的卷积层，以充分利用两个层的优势，新的目标检测模块如图 5 所示。

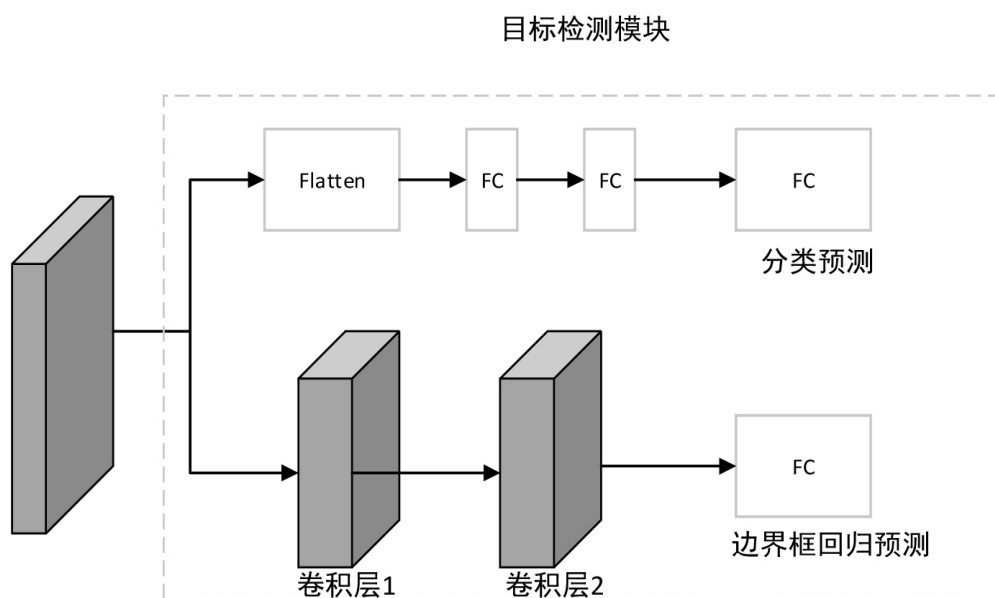


Figure 5. New target detection module
图 5. 新的目标检测模块

改进的网络结构与 Faster RCNN 的一样是两阶段的检测网络，使用 RPN 网络获取候选框，有高精度的检测性能，可解决多尺度、小目标问题。改进的网络结构在预测层与原 Faster RCNN 网络不同，对分类和定位两个任务分别进行处理，所以有更好的检测效果。改进网络整体框架如图 6 所示，在 ROIpooling 层后，网络导入两个分支，一个分支采用卷积进行特征提取并进行边界框回归预测，一个分支采用全连接层进行特征提取并进行分类预测。

5. 矿工出入三线检测算法

图片通过改进 Faster RCNN 后获取安全帽目标，为实现人员统计，借鉴文献[13]的基于计算机视觉的目标计数方法和文献[14]的基于检测的方法和判别式跟踪技术，本文设计了三线检测算法，如图 7 所示。算法设定检测目标顺序越过第一线后再越过第二线(进入矿井)或者越过第三线后再越过第二线(离开矿井)才被计数。越过第二线时如果目标对角线两点的 y 坐标比上一帧的两点的 y 坐标值大，则判定目标进入矿井；如果目标对角线两点的 y 坐标比上一帧两点的 y 坐标值小，判定目标离开矿井。三线检测算法描述如算法 1 所示。算法流程图如图 8 所示。

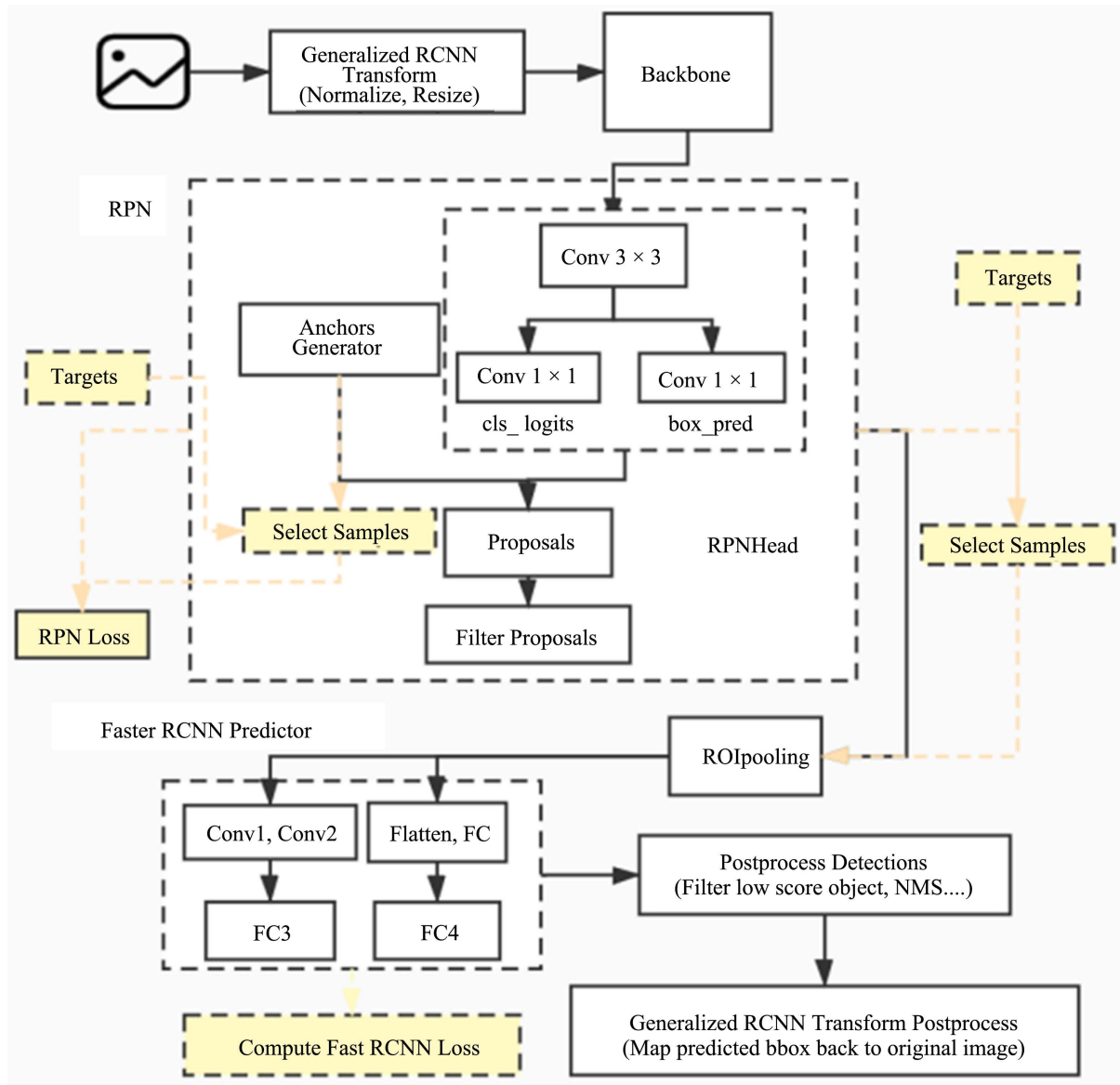


Figure 6. Safety helmet detection network based on improved faster RCNN

图 6. 改进 Faster RCNN 的安全帽检测网络

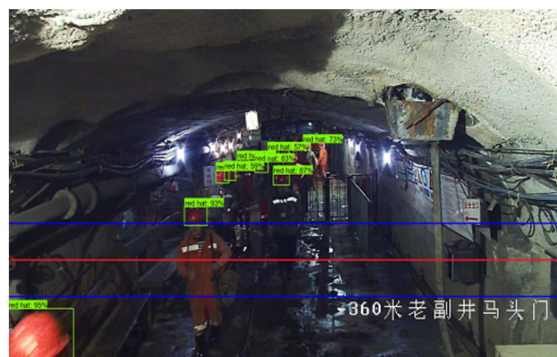


Figure 7. Three line detection of real image

图 7. 三线检测真实图像

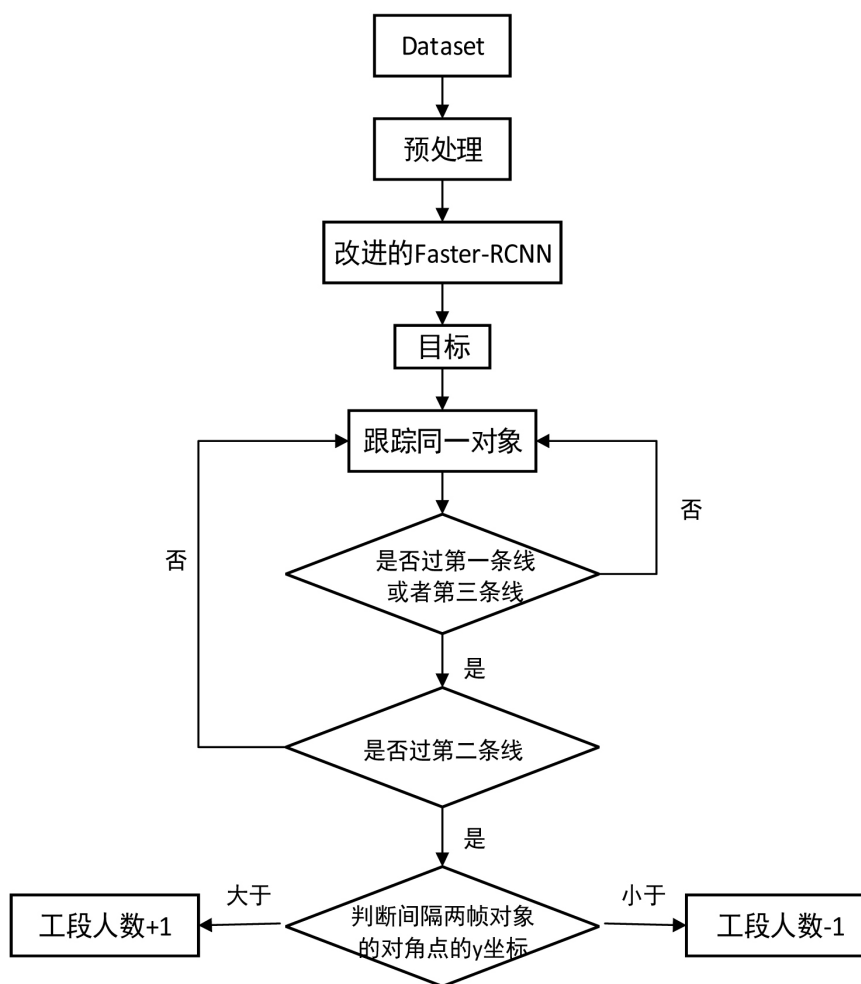


Figure 8. Flow chart of three line detection algorithm

图 8. 三线检测算法流程图

算法 1: 矿工出入三线检测算法

输入: pt1 目标左上角坐标, pt2 目标右下角坐标 //包括 x 坐标和 y 坐标

输出: Num 该工段人数

1. **function** PeopleNum(pt1,pt2)
2. **if** pt2.y > tFrame[i][1].y \wedge pt1.y > tFrame[i][0].y \wedge pt2.y > SecondLine \wedge pt1.y < SecondLine \wedge tMark[i] == false **then**//下井的矿工人数统计
3. tMark[i] \leftarrow true;
4. tUpOrDown[i] \leftarrow true;
5. TotalNum++;
6. Num++;
7. NumDown++;
8. **if** pt1.y < tFrame[i - 1][0].y \wedge pt1.y <= 870 \wedge pt2.y < tFrame[i - 1][1].y \wedge tMark[i - 1] == false **then**
//上井的矿工人数统计
9. tMark[i] \leftarrow true;

```

10.   tUpOrDown[i] ← false;
11.   TotalNum--;
12.   Num--;
13.   NumuP++;
14.   isNew ← false;
15.   if isNew && pt1.y <= FirstLine && pt2.y>FirstLine && pt2.y <= ThirdLine then
                                     //过第一线或者第三线进行标记
16.   int i = 0;
17.   while i < 200
18.   if ! fMark[i]&&! tMark [i]
19.   fMark[i] = true;
20.       break;
21.       i++;
22.   return Num
23.   end function

```

6. 实验与结果分析

6.1. 评价指标

本文采用的评价指标[15]是平均查全率(Average Recall, AR)、平均准确率(Average Precision, AP)和平均精度均值(mean Average Precision, mAP)查全率和平均精度均值越高算法效果越好。查全率和准确率定义如式 1 所示。

$$P = \frac{TP}{TP + FP} \quad (1)$$

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

其中, TP (True Positive)是被判别为正样本的正样本数量, FN (False negative)是被误判为负样本的正样本数量。FP (False Positive)是被误判为正样本的负样本数量。

AP 计算如式 2 所示。

$$AP = \int_0^1 p(r) d(r) \quad (2)$$

平均精度均值 mAP 是由 Precision-recall 曲线与坐标轴包围区域的面积。c 为计算的组数, mAP 计算如式 3 所示。

$$mAP = \frac{\sum_{c=1}^c AP(c)}{c} \quad (3)$$

6.2. 实验结果与分析

使用井下矿工数据集训练本文提出的改进的 Faster RCNN 算法。为了证明方法的有效性和合理性, 本文对实际中采集的井下行人数据集进行了训练、测试和验证。井下矿工行走数据集均在提升机罐笼进出口提取, 矿工均保证佩戴安全帽。井下矿工行走数据集共 200 张图片, 其中训练集数据 140 张, 测试

集数据 30 张, 验证集数据 30 张。训练时初始学习率设置为 0.005, 在训练 10 个批次、15 个批次和 20 个批次后分别衰减, 在 25 个批次之后停止训练。图 9 展示了井下矿工数据集的损失曲线。从损失曲线中可以看出在开始时迅速下降, 并随着学习率的降低, 损失曲线缓慢且稳步地下降并在第 3 个批次时, 训练损失曲线损失值稳定在 0.2 左右。该数据说明本文改进的 Faster RCNN 收敛速度快, 且收敛效果好。损失在 0.2~0.3 之间小幅震荡。最终得到的模型的评价指标 mAP 曲线如图 10 所示。

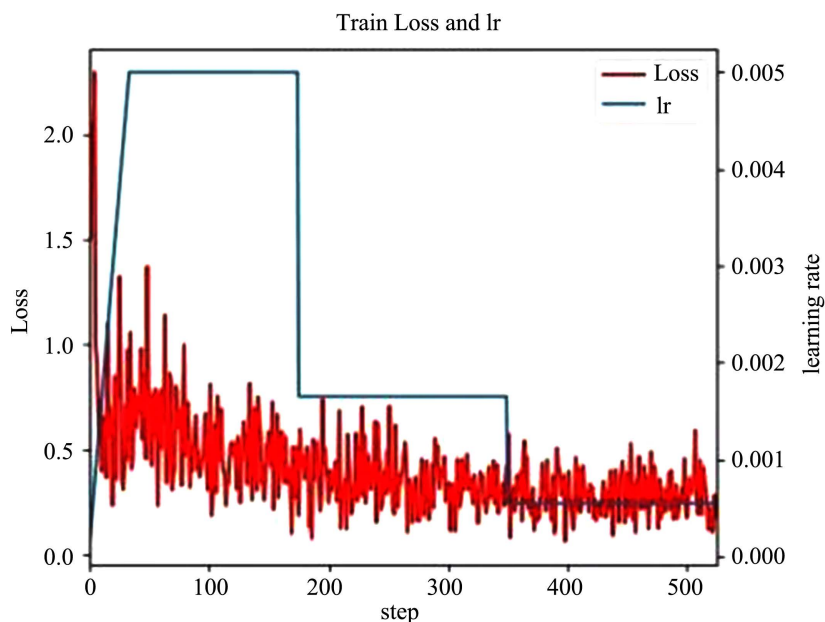


Figure 9. Loss function and learning rate curve

图 9. 损失函数和学习率曲线

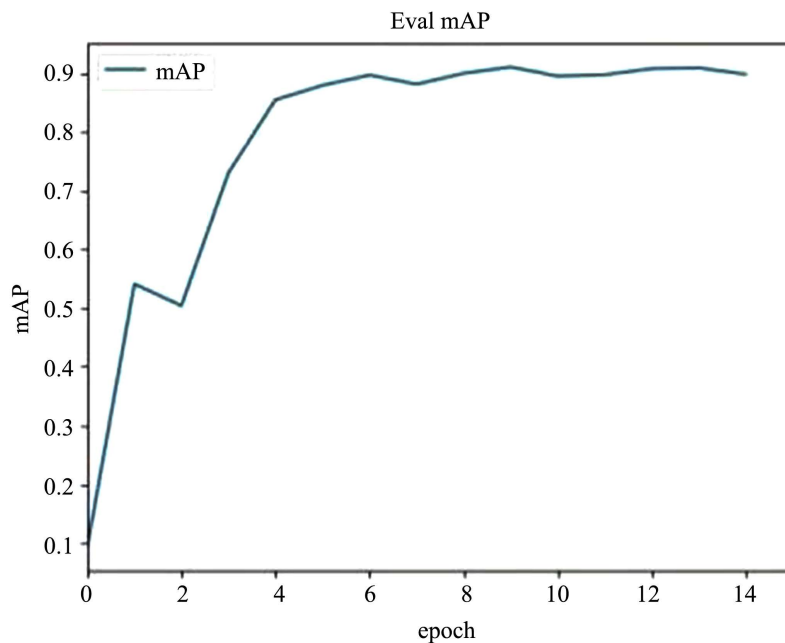


Figure 10. mAP curve

图 10. mAP 曲线

为了验证改进算法的可行性和先进性,改进的 Faster RCNN 算法与传统 Faster RCNN 算法在井下矿工行走数据集的验证集上进行了对比,结果如表 1 所示。

Table 1. Algorithm comparison results display
表 1. 算法对比结果展示

IoU = 0.5	IoU = 0.75	IoU = 0.5:0.95	Area = small	Area = medium	Area = large	Area = all	maxDets = 1	maxDets = 10	maxDets = 100	传统 Faster RCNN AP%	改进 Faster RCNN AP%	传统 Faster RCNN AR%	改进 Faster RCNN AR%
		√				√			√	36.2	44.8	43.4	52.4
√						√			√	88.7	91.8	-	-
	√					√			√	21.7	36.4	-	-
		√	√						√	32.6	41.8	40.3	49.2
		√		√					√	50.7	57.1	40.3	49.2
		√			√				√	63.8	65.1	65.0	72.5
		√				√	√			-	-	8.3	8.4
		√				√		√		-	-	39.8	47.0

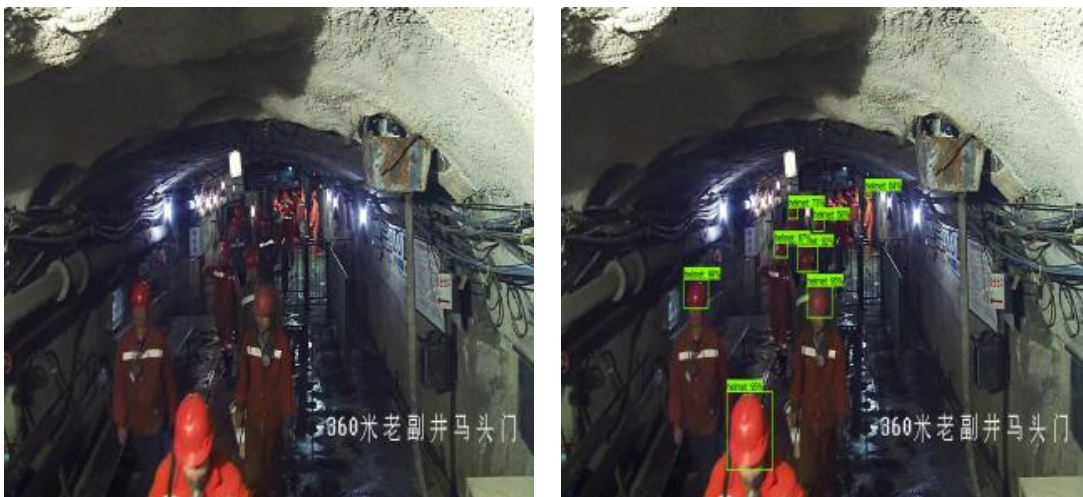
表中 IoU 表示可判定为同一对象的交并比阈值,area 表示目标尺度,maxDets 表示每幅图像最大检测次数。AP 表示平均准确率,AR 表示平均查全率。与传统 Faster RCNN 目标检测算法相比,改进的 Faster RCNN 算法对所有安全帽样本不同比例大小的目标平均查准率提高了 8.8%,平均查全率提高了 5.4%。针对不同尺度安全帽目标的检测结果如下:

对于小尺度目标,改进的算法平均查准率提高 9.2%,平均查全率提高 8.9%;

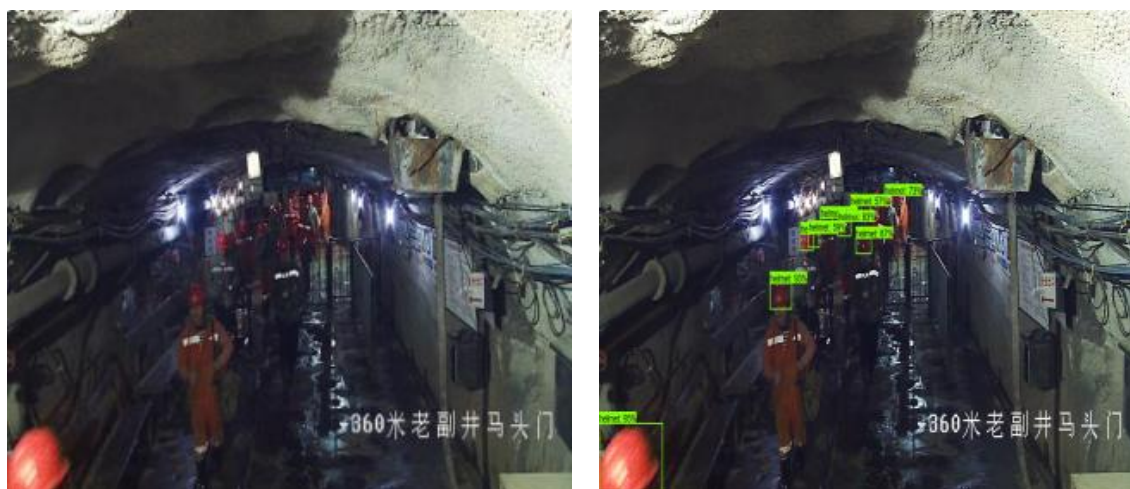
对于中尺度目标,改进的算法平均查准率提高 6.4%,平均查全率提高 9.4%;

对于大尺度目标,改进的算法平均查准率提高 1.3%,平均查全率提高 7.5%。

更高的平均查准率和平均查全率可以减少矿工检测错漏。选取其中具有代表性的图像,用改进的 Faster RCNN 算法进行检测,检测结果如图 11 所示。



(a) 真实井下图像 1



(b) 真实井下图像 2



(c) 真实井下图像 3



(d) 真实井下图像 4

Figure 11. Detection effect of miner safety helmet
图 11. 矿工安全帽检测效果

由图 11 可以看出,改进的 Faster RCNN 算法在实际工程中对井下矿工安全帽的识别效果较好,而且对有遮挡的安全帽也可以准确识别出,配合本文提出的出入三线检测算法后可实现人数统计。

7. 结论

为解决井下矿工人数统计问题,从矿工安全帽检测出发,本文通过改进分类和定位预测网络得到改进的 Faster RCNN 框架。在真实井下矿工图像数据集上对比实验表明:与传统的 Faster RCNN 相比,改进的 Faster RCNN 框架检测算法平均准确率提高了 8.8%,平均查全率提高了 5.4%,即使在光线很差时出现漏检的情况,通过设计三线检测算法后可确保人数统计正确。下一步计划对算法进行进一步改进,加入人脸识别的分类器,在实现人数统计的同时,完成考勤功能。

基金项目

广东省重点领域研发计划 2021B0101200002、2021B0101200002、2019B01018001。

参考文献

- [1] 王周春. 基于机器学习的目标探测与识别技术研究[D]: [硕士学位论文]. 北京: 中国科学院大学(中国科学院上海技术物理研究所), 2020.
- [2] 齐燕舞, 朱杰. 基于 Haar-like 特征多分类器集成的行人检测[J]. 信息技术, 2017(8): 129-131.
- [3] 任小康, 陈鸿享. 基于 HOG+SVM 实现行人检测[J]. 科学技术创新, 2021(6): 73-74.
- [4] 芮挺, 费建超, 周遊, 等. 基于深度卷积神经网络的行人检测[J]. 计算机工程与应用, 2016, 52(13): 162-166.
- [5] 吴冬梅, 王慧, 李佳. 基于改进 Faster RCNN 的安全帽检测及身份识别[J]. 信息技术与信息化, 2020(1): 17-20.
- [6] 曹燕, 宋正伟, 周楠. 一种基于等级 SSD 的筑工人安全帽佩戴检测方法[J]. 工业控制计算机, 2021, 34(1): 55-56.
- [7] 肖体刚, 蔡乐才, 高祥, 等. 改进 YOLOv3 的安全帽佩戴检测方法[J]. 计算机工程与应用, 2021, 57(12): 216-223.
- [8] Wu, Y., Chen, Y.P., et al. (2020) Rethinking Classification and Localization for Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 14-19 June 2020, 10186-10195. <https://doi.org/10.1109/CVPR42600.2020.01020>
- [9] Qin, Y.Y., He, S.Y., Zhao, Y. and Gong, Y.Z. (2016) RoI Pooling Based Fast Multi-Domain Convolutional Neural Networks for Visual Tracking. *Proceedings of 2016 2nd International Conference on Artificial Intelligence and Industrial Engineering (AIIE2016)*, Beijing, 20-21 November 2016, 198-202. <https://doi.org/10.2991/aiie-16.2016.46>
- [10] He, K., Zhang, X., Ren, S., Sun, J., et al. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 26 June-1 July 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [11] 王亚南. 基于 RPN 网络的行人检测方法研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2020.
- [12] 孙世强, 左海维, 赵露婷. 联合特征相似性度量和交并比的检测框优选研究[J]. 电脑知识与技术, 2019, 15(29): 190-193.
- [13] 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.
- [14] 尹宏鹏, 陈波, 柴毅, 等. 基于视觉的目标检测与跟踪综述[J]. 自动化学报, 2016, 42(10): 1466-1489.
- [15] 鲍捷, 吕春, 罗庚, 等. 一种目标检测算法的性能评估方法及系统[P]. 中国专利, CN110991485A. 2020-04-10.