

引文格式：

张传深, 徐升, 胡佳, 等. 基于巡逻无人机的轻量型安全帽佩戴检测方法与应用 [J]. 集成技术, 2023, 12(4): 18-31.

Zhang CS, Xu S, Hu J, et al. Research on safety helmet recognition method and application using patrol unmanned aerial vehicle [J]. Journal of Integration Technology, 2023, 12(4): 18-31.

基于巡逻无人机的轻量型安全帽佩戴检测方法与应用

张传深^{1,2} 徐升^{3*} 胡佳⁴ 王强¹

¹(山东中科先进技术有限公司 济南 250102)

²(中国移动通信集团山东有限公司德州分公司 德州 253521)

³(中国科学院深圳先进技术研究院 深圳 518055)

⁴(浪潮集团有限公司 济南 250101)

摘要 目前, 安全帽检测系统主要使用固定摄像头, 无法实现全区域检测, 而基于深度学习的检测算法结构复杂、计算成本高, 无法满足移动端和嵌入式设备的部署要求。针对上述问题, 该文提出一种基于无人机的安全帽轻量型视觉检测算法。系统通过无人机平台搭载的相机对施工现场进行图像采集, 并无线传输至后台计算机进行处理, 检测算法基于 YOLOv5s 框架进行了轻量化改进。针对无人机采集影像中目标占比较小的问题, 该文采用了多尺度检测、图像预处理、正负样本不均衡等方法, 对 YOLOv5s 目标检测算法进行针对性改进。测试结果表明, 与原模型相比, 轻量型目标检测模型的平均精度均值仅下降了 1.72%, 但在同一 CPU 上的推理速度提升了 1 倍, 浮点计算量由原来的每秒 165 亿次压缩至每秒 34 亿次, 模型大小约为原模型的 1/10。

关键词 轻量型检测算法; 安全帽检测; 无人机; 自动巡航

中图分类号 TP 39 文献标志码 A doi: 10.12146/j.issn.2095-3135.20220720001

Research on Safety Helmet Recognition Method and Application Using Patrol Unmanned Aerial Vehicle

ZHANG Chuanshen^{1,2} XU Sheng^{3*} HU Jia⁴ WANG Qiang¹

¹(Shandong Institute of Advanced Technology, Chinese Academy of Sciences, Jinan 250102, China)

²(China Mobile Shandong Co., Ltd., Dezhou Branch, Dezhou 253521, China)

³(Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

⁴(Inspur Group Co., Ltd., Jinan 250101, China)

*Corresponding Author: sheng.xu@siat.ac.cn

Abstract The existing helmet detection system mainly uses a fixed camera, it cannot achieve full-area

收稿日期: 2022-07-20 修回日期: 2023-02-22

基金项目: 山东省重点研发计划项目(2021CXGC011304); 深圳市科技计划资助项目(JCYJ20210324102401005)

作者简介: 张传深, 硕士, 研究方向为嵌入式与人工智能; 徐升(通讯作者), 副研究员, 硕士研究生导师, 研究方向为目标定位、信号处理、智能机器人, E-mail: sheng.xu@siat.ac.cn; 胡佳, 硕士, 研究方向为无人机组合导航; 王强, 硕士, 研究方向为机器人开发、多传感器融合。

detection, and the previous detection algorithms based on deep learning have complex structures and high computational costs, which cannot meet the requirements of using mobile vehicles and embedded devices. In this paper, a lightweight helmet detection algorithm scheme based on unmanned aerial vehicle is proposed. The drone is loaded with camera to collect images of the construction site, and the image data is transferred to the computer via wireless communication. Based on the YOLOv5s target detection algorithm, a lightweight detection algorithm is investigated. To improve the detection more efficient, the YOLOv5s target detection algorithm is improved in terms of multi-scale detection, image preprocessing, unbalanced positive and negative samples, and inference speed. This design scheme combines deep learning and unmanned aerial vehicle technology, not only to realize real-time automatic detection of helmet wearing, but also can realize the full-area helmet detection of the construction site. Real experiments show that, the lightweight target detection model is only 1.72% lower than the mean average precision of the original model. The inference speed on the same CPU can be doubled, and the floating-point calculation is reduced from 16.5 billion to 3.4 billion times per second. The model size is almost 1/10 of the original size.

Keywords lightweight detection algorithm; safety helmet detection; unmanned aerial vehicle; automatic patrol

Funding This work is supported by Key Research and Development Program of Shandong Province (2021CXGC011304), and Shenzhen Science and Technology Program (JCYJ20210324102401005)

1 引言

建筑工地中, 佩戴安全帽有较高的防护作用。但在实际工作中, 很多工人由于施工原因不愿佩戴安全帽, 导致安全事故时有发生。目前, 针对工人不佩戴安全帽的行为, 各施工工地多采用人工巡查结合摄像头监控的方式进行检查, 这不仅造成了人力、物力的浪费, 而且效率低下。

随着科学技术的进步, 众多学者对如何自动实时地检测工人佩戴安全帽展开了广泛的研究。朱富云等^[1]通过人体几何特征对头部进行定位, 然后使用模板匹配的方法进行安全帽匹配。Waranusast 等^[2]使用 K -近邻 (K -nearest neighbor, K -NN) 分类器从区域属性中提取特征并进行分类, 完成摩托车骑手头部识别, 然后根据投影分析, 对识别出的骑手头部进行计数和分割。进而, 基于头部区域分割的特征, 利用 K -NN 分类

器将头部分为戴安全帽或不戴安全帽。Rubaiyat 等^[3]将图像的频域信息与流行的人体检测算法定向梯度直方图相结合, 用于识别建筑工人, 然后基于颜色和圆霍夫变换特征提取技术, 检测建筑工人的头盔使用情况。此外, 对于除安全帽之外的各类人体佩戴物或运动人员等目标, 也有大量学者提出了基于图像处理的各类方法^[4-5]。上述研究虽然取得了一定成果, 但仍存在特征提取困难、泛化性能差等问题, 随着深度学习技术的发展, 为目标检测提供了许多新方法^[6-7]。

基于深度学习的目标检测方法一般包括双阶段检测的区域卷积神经网络 (region convolutional neural network, RCNN)、空间金字塔池化网络、快速卷积神经网络、超快卷积神经网络 (Faster RCNN)、遮掩卷积神经网络和单阶段检测的一次观测法 (You Only Look Once, YOLO) 系列和单观测多窗口检测。双阶段检测的检测精度高,

但因网络结构限制,检测速度较慢;单阶段检测是基于回归的检测算法,虽然检测精度略逊于双阶段检测,但其检测速度非常快。目前,有大量研究学者将上述基于深度学习的目标检测方法用于安全帽的检测。张勇等^[8]使用改进的YOLOv3对安全帽进行检测,准确率为96.5%,检测速度为59.6 fps。孙国栋等^[9]通过融合自注意力机制改进Faster RCNN,与传统的Faster RCNN相比,其平均精度均值(mean average precision, mAP)提高了4%。张明媛等^[10]使用Faster RCNN目标检测网络,实时监测工人的安全帽佩戴状况,该模型的平均精度为90.91%,召回率为89.19%。王忠玉^[11]利用一种融合深度卷积神经网络固有的特征金字塔多层特征进行分层预测的方法,并基于Faster RCNN进行改进,设计了一种基于深度学习的安全帽佩戴检测算法。徐守坤等^[12]基于原始Faster RCNN,利用多尺度训练等方法增强网络检测不同尺寸目标的鲁棒性,并对正负样本不均衡进行处理。夏鼎等^[13]提出包含部件特征增强、部件特征融合及人体特征增强模块的深度学习方法,以改善人体部位及行为识别准确性。王雨生等^[14]首先通过肤色识别和头部识别进行交叉验证确定头部区域,然后使用YOLOv4目标检测网络识别安全帽,最后通过安全帽区域与头部区域的位置关系,判断安全帽的佩戴情况。韩锟等^[15]对YOLOv3的网络结构、损失函数及先验框尺寸进行改进,提出YOLOv3-C模型,该模型的检测性能较YOLOv3有大幅提升。

上述基于深度学习的安全帽检测方法,虽然检测速度快、准确度高,但是结构较为复杂、计算量过大,难以满足实际应用的需求。针对上述问题,本研究提出一种轻量型安全帽目标检测算法。此外,无人机作为一种灵活的载具,可以高效实现对地观察,从而推动各类监测应用的发展^[16-17]。本文应用自主研发的巡逻无人机^[18]进行

施工工地的图像采集,有效解决了固定摄像头视野范围受限的问题,并且可以替代人工巡逻,节省人力成本的同时,提高了安全帽的检测效率,保证了工人的安全。

2 YOLO 目标检测原理

YOLOv5是YOLO系列最新的检测网络,其图像推理速度最快可达到每秒140帧(140 fps),满足实时检测需求。另外,与YOLOv4相比,YOLOv5结构更为小巧,权重文件最小为14 MB。虽然YOLOv5模型在GPU上的检测速度较快,模型体积较小,但对于移动端或嵌入式设备,仍不能满足其计算量和模型所占内存,无法进行部署与应用。针对该问题,本文从减小算法深度、重构算法结构两个方面,对YOLOv5s目标检测算法进行模型轻量化设计,在保证检测精度微小下降的前提下,减少模型的参数和计算量,加快模型的推理速度。针对无人机采集图像中目标占比较小,识别准确率差的问题,本文对网络的空间金字塔池(spatial pyramid pooling, SPP)结构进行改进。此外,为提高目标检测模型对安全帽检测的准确性,还使用CIoU损失替换GIoU损失。

YOLO算法摒弃了串口窗口滑动的思想,直接将原始图像分割成 $S \times S$ 的小方格,并通过卷积生成 $S \times S$ 特征图,根据滑动窗口思想,可认为特征图的每个元素都对应原始图像的一个小方格。然后,用特征图的每个元素预测中心在该格子内的目标,即YOLO算法最基本的思想^[19]。YOLOv5目标检测网络通过图1中各模块一系列的卷积、归一化、激活、池化、上采样等操作,将输入尺寸为 $[640, 640]$ 的图片的特征提取至 20×20 、 40×40 、 80×80 ,3个不同尺寸的特征图,以预测不同大小的目标。特征图通道为 $3 \times (1 + 4 + C)$ 。其中,4为先验锚框坐标相对于预测框坐标的偏移量 (t_x, t_y, t_w, t_h) 。C为每个类别

对应的概率, 本研究中 C 为 2.3 为每个网格预测先验框(anchor box)的个数, YOLOv5 会产生 9 种不同大小的先验框, 对应不同尺寸的特征, 为每种尺寸的特征图产生 3 个先验框。1 为边界框的置信度, 其包含两个方面: (1) 该边界框含有目标的可能性大小, 记为 $p_r(\text{object})$, 当该边界框为背景时(即不包含目标), $p_r(\text{object})=0$, 当该边界框包含目标时, $p_r(\text{object})=1$ 。(2) 该边界框的准确度, 可以用预测框与实际框(ground truth)的交并比(intersection over union, IOU)进行表征, 记为 $\text{IOU}_{\text{pred}}^{\text{truth}}$ 。因此, 置信度的定义可如公式(1)所示。

$$\text{IOU}_{\text{pred}}^{\text{truth}} \times p_r(\text{object}) \quad (1)$$

上述各概率值是在各个边界框置信度下的条件概率, 即 $P_r(\text{class}_i|\text{object})$ 。先验框类别置信度为类别概率与边界框置信度的乘积, 其既预测了锚框属于某一类的类别, 又包含了锚框的准确度, 计算公式如公式(2)所示。

$$P_r(\text{class}_i|\text{object}) \times \text{IOU}_{\text{pred}}^{\text{truth}} \times p_r(\text{object}) \quad (2)$$

对于尺寸为[640,640]的图片, 经 YOLOv5 后输出 $(20 \times 20 + 40 \times 40 + 80 \times 80) \times 3 = 25\ 200$ 个先验框。这些先验框存在大量冗余, 需要通过非极大值抑制的方法进行筛选, 即首先从所有的先验框中找到置信度最大者, 逐个计算其与剩余框的交并比, 若此比值大于设定阈值, 代表重合度过高, 那么将该框剔除, 然后对剩余的检测框重复上述过程, 直至处理完所有的检测框。

经过上述所有步骤后, 进行反复迭代计算预测框与真实框的 LOSS。当 LOSS 值最小时, 输出值为最优的坐标偏移值、置信度得分和类别概率, 然后使用最优的坐标偏移值调整先验框的位置, 使之与真实框无限接近。YOLOv5 的损失函数包括坐标偏移量损失、置信度损失和类别概率损失。置信度和类别概率损失采用的是二进制交叉熵损失函数(公式(3)), 坐标偏移量损失采用的是 GIoU 损失(公式(4))。

$$\text{LOSS} = -(y \cdot \log(\hat{y}) + (1-y) \cdot \log(1-\hat{y})) \quad (3)$$

其中, \hat{y} 为模型预测样本是正例的概率; y 为样本标签, 若样本属于正例, 那么 y 取值为 1, 否则取值为 0。

$$L_{\text{GIoU}} = 1 - \text{IOU}(A, B) + |C - A \cup B| / |C| \quad (4)$$

其中, A 为预测框; B 为真实框; C 为 A 和 B 的最小包围框。

3 训练数据集及训练环境

训练数据集是基于北京旷视科技有限公司发布的 CrowdHuman、网络搜集和手机拍摄创建。CrowdHuman 数据集是旷世科技发布的用于行人检测的数据集, 图片数据大多来自 Google 搜索, 共计 19 470 张, 每张图片中的人类都用头部边界框和人体全身体边界框进行标注。对于 CrowdHuman 数据集, 首先将其标注文件格式转化为 YOLO 格式, 即用文本文件保存标注信息, 文件第 1 列数字为目标类别, 第 2~5 列为标注目标的位置信息, 且数值全部归一化处理, 将人体全身体边界框的标注信息删除, 仅保留人类头部的标注信息。在 CrowdHuman 数据集中, 人类实例多为西方人, 且图片多经过美化加工, 考虑到数据集分布的均衡性, 本实验使用网络搜集和手机拍摄的方法获取 3 000 张东方人图片, 不进行处理, 并在数据集中加入少量不包含目标的背景图片, 便于模型的学习, 降低误检率。

本文采用半自动的方式对搜集和拍摄的图片进行标注, 首先在 CrowdHuman 数据集上使用 YOLOv5s 模型进行预训练, 然后在收集与拍摄的图片上进行推理, 记录并保存推理结果, 使用 labellmg 标注工具进行微调, 该方法大大缩短了数据集标注的时间。

本文研究方案在原有 CrowdHuman 数据集基础上创建包含 22 470 张图片的数据集, 以满足目标检测网络的训练要求。表 1 为数据集中训练样

本的具体分类情况，数据集被分为训练集、验证集和测试集。另外，在训练过程中，每次迭代结束后均会对当前训练结果进行反馈，以防止过拟合情况。

网络训练实验环境使用 Ubuntu18.04 操作系统，选用 Pytorch 框架搭建目标检测网络，使用 Tesla V100 的显卡进行加速运算，具体实验配置如表 2 所示。

表 1 安全帽检测数据集划分

Table 1 The data sets of the safety helmet detection

数据集类型	数量 (张)
总数据集	22 470
训练集	13 482
验证集	6 741
测试集	2 247

表 2 实验环境参数

Table 2 Details of the experiment environment

配置参数	
CPU	Intel(R) Xeon(R) Gold 6254 CPU @ 3.10 GHz
GPU	Tesla V100
系统环境	Ubuntu18.04
语言	Python3.8
加速环境	CUDA10.2

4 轻量化设计及性能优化

4.1 重新设计深度系数

本研究选择不同的深度系数进行试验，分析

深度系数与 mAP、浮点数计算量和模型参数之间的关系，部分实验结果如表 3 所示。

由表 3 可知，与深度系数为 0.50 (YOLOv5s) 的模型相比，当深度系数重新设计为 0.25 (YOLOv5n) 时，虽然模型参数和浮点数计算量大幅下降，但是主要性能评价指标 mAP 也出现较大幅度的下降。当深度系数为 0.40 时，模型参数和浮点数计算量有大幅下降的同时，mAP 并未出现较大幅度的下降。因此，本实验将 YOLOv5 的深度系数重新设计为 0.40，后续对特征提取网络、推理网络的修改和对模型整体的稀疏化训练、模型剪枝均基于此模型。

4.2 重构特征提取网络

针对特征提取网络在进行特征提取时存在大量冗余的问题，本实验以 YOLOv5-0.40 检测网络为基础，使用 GhostNet 模型中“幻影卷积 (GhostConv)”设计的 GhostBottleneck 模块^[20]和 MobileNetV3 网络模型中的“倒残差模块 (InvertedResidual)”^[21]，分别对特征提取网络进行重构，并通过实验进行性能对比，选择最优重构方案。

GhostNet 是华为诺亚方舟实验室提出的新型端侧神经网络架构，其提出者认为提取的特征中一般含有大量冗余特征图，以保证对输入数据有详细的特征理解。且在特征提取的过程中，仅部分特征图通过卷积操作得到，对于冗余的特征图可通过更简单的线性变换得到。因此，GhostNet

表 3 重新设计模型深度系数实验结果

Table 3 The experiment results of with the re-designed model and parameters

	平均精度 mAP (%)	每秒处理帧数 FPS (帧/s)	精准率 (%)	召回率 (%)	每秒十亿次浮点运算数 GFLOPS (十亿次/s)	模型参数存储大小 (MB)
YOLOv5s	79.88	6	86.46	72.16	16.5	13.6
YOLOv5-0.45	79.67	6	86.51	71.89	13.8	11.3
YOLOv5-0.40	79.53	8	86.18	71.50	11.1	9.16
YOLOv5-0.35	78.21	8	86.52	70.07	8.6	7.13
YOLOv5-0.30	77.07	8	86.07	68.98	6.5	5.41
YOLOv5n	75.70	9	85.16	67.76	4.2	3.65

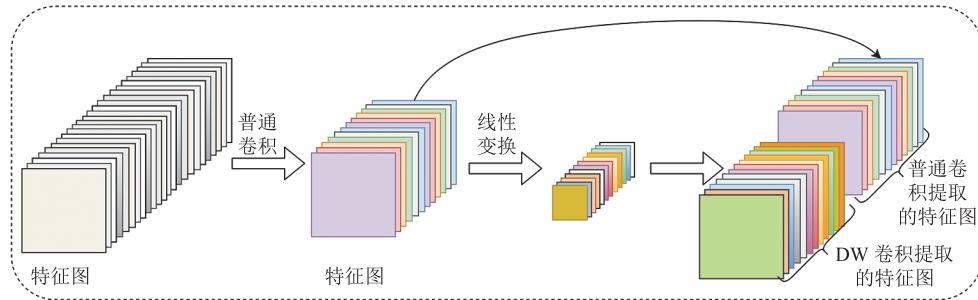


图 1 GhostConv 结构图

Fig. 1 Diagram of the GhostConv

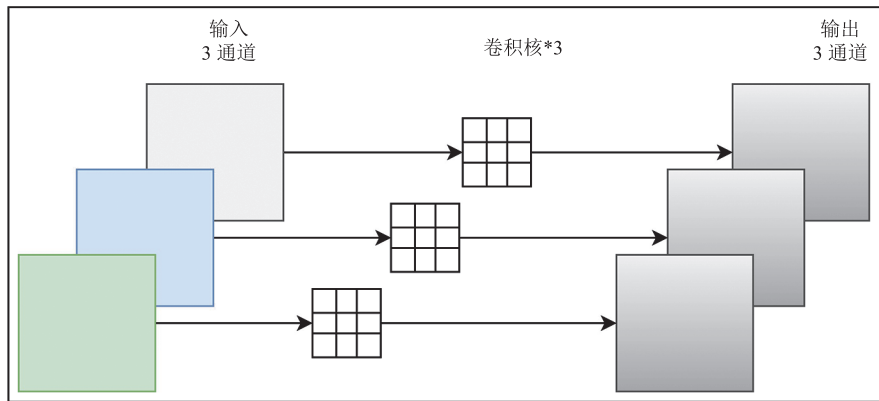


图 2 深度卷积网络结构

Fig. 2 The structure of the depthwise convolution neural network

的作者提出了 GhostConv, 其网络结构如图 1 所示。首先通过普通卷积操作提取一部分特征图, 然后对特征图进行线性变换, 最后经过普通卷积和线性变换得到两部分特征图, 并在通道维度上进行组合, 得到最终的特征图。

MobileNetV3 是 Google 提出的轻量化网络模型, 相较于 MobileNetV2 加入了注意力机制, 更新了激活函数, 使用 $h\text{-swish}[x]$ 函数作为网络的激活函数, 重新对耗时层结构进行设计, 将第 1 层卷积的卷积核个数减少至 16 个。在分类任务上, 与 MobileNetV2 相比, MobileNetV3-Small 的分类精度提高了 3.2%, 推理时间降低了 15%; MobileNetV3-Large 分类精度提高 4.6%, 推理时间降低 5%。此外, 以 MobileNetV3-Large 和 MobileNetV2 作为主干网络的两种目标检测算法, 在 COCO 检测数据集上分别运行。

结果表明, 两种算法的检测精度相同, 但是以 MobileNetV3-Large 为主干网络的算法推理速度较 MobileNetV2 加快了 25%, 同时在分割算法上也更优^[22]。

使用 MobileNetV3 主模块 (InvertedResidual) 重构特征提取网络的目标检测模型为 YOLOv5s-MobileNetV3, 其中, 深度卷积 (depthwise convolution, DWConv) 的网络结构如图 2 所示, 重构后的特征提取网络结构图如图 3 所示。

使用 GhostConv 设计 GhostBottleneck 模块, 网络结构如图 4 所示, 采用两种方案基于其重构特征提取网络: (1) 方案一如图 5(a) 所示, 将其命名为 YOLOv5s-Ghost, 该方案使用 GhostBottleneck 模块替换原始特征提取网络中大部分的 C3 模块, 输入端卷积层保持不变, 将其余卷积层改为图 6 所示结构的深度可分卷积

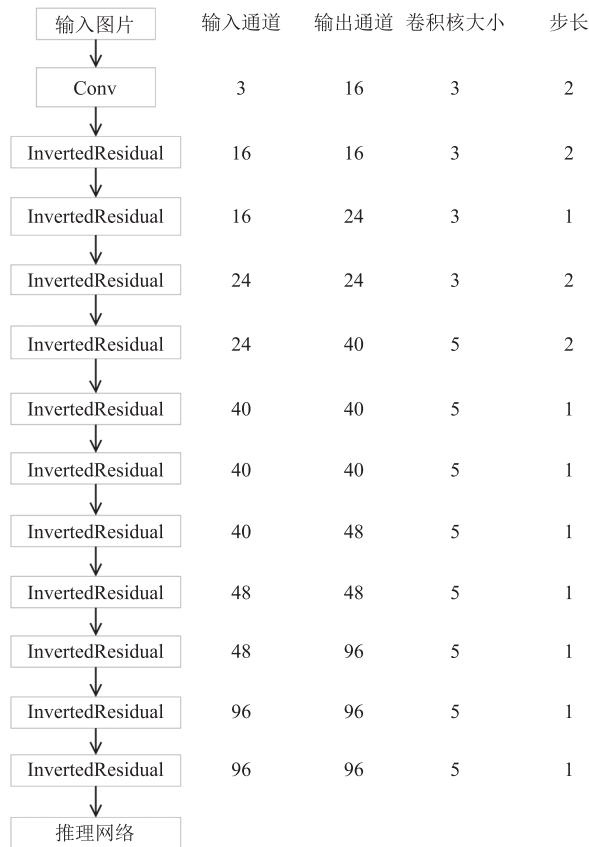
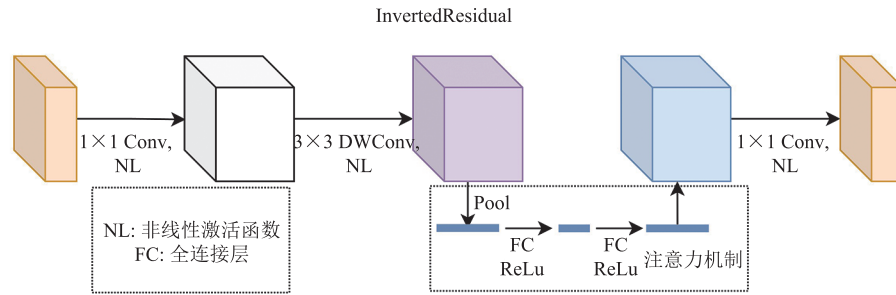


图 3 YOLOv5s-MobileNetV3 特征提取网络结构

Fig. 3 The feature extraction structure of the YOLOv5s-MobileNetV3

(depthwise separable convolution, DSConv), 以减小模型的参数; (2) 方案二如图 5 (b) 所示, 将其命名为 YOLOv5s-Ghost-SE, 在该方案中, GhostBottleneck 模块的部署与方案一大致相同, 但此方案没有使用深度可分卷积, 而是将原始特征提取网络结构中间部分的两个卷积层使用 GhostBottleneck 模块代替, 并加入注意力机制。

表 4 为 YOLOv5s-MobileNetV3、YOLOv5s-

Ghost 和 YOLOv5s-Ghost-SE 在人类头部数据集上经 100 次迭代后的各项评估指标。由表 4 可知, InvertedResidual 模块和 GhostBottleneck 模块重构特征提取网络都可使得模型参数和浮点数计算量大幅减小, 但与原始模型相比, YOLOv5s-MobileNetV3 的 FPS 较小, 即推理速度较慢, YOLOv5s-Ghost-SE 的 mAP 下降幅度较大, YOLOv5s-Ghost 的模型参数和计算量大幅下降,

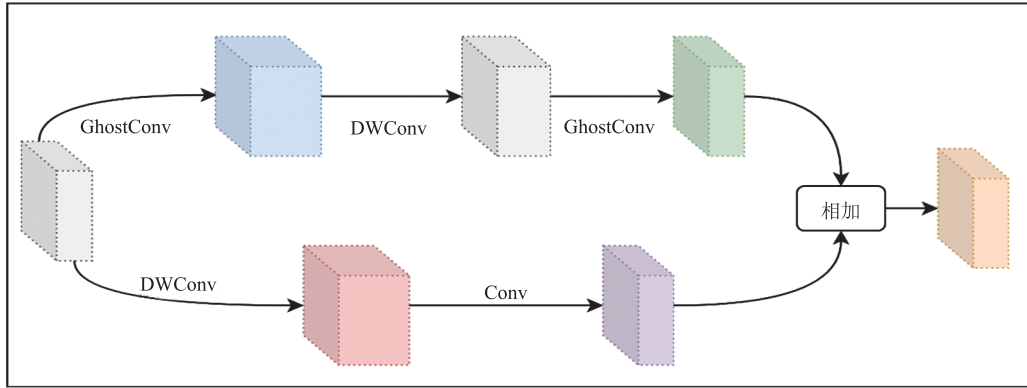
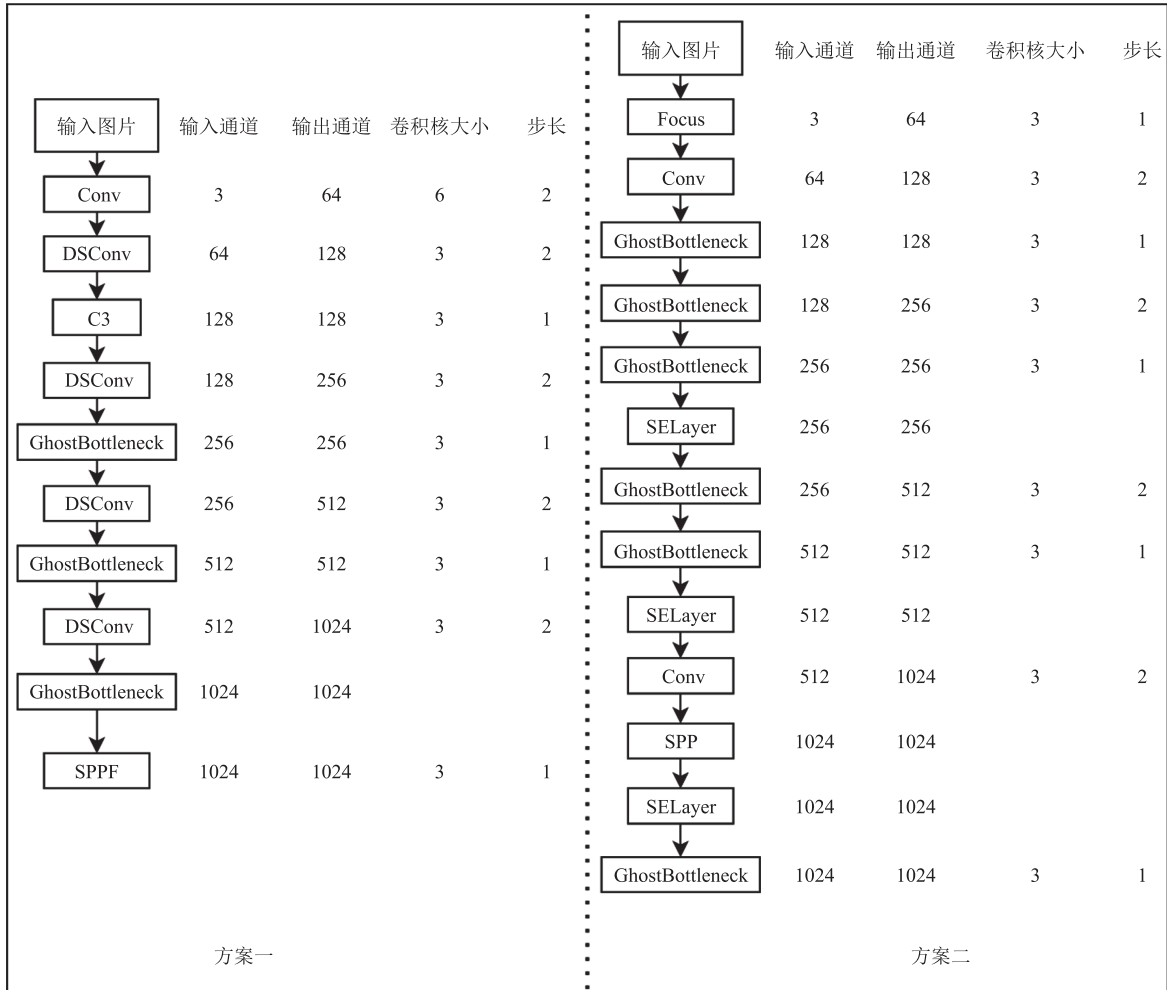


图 4 GhostBottleneck 模块网络结构

Fig. 4 Structure of the GhostBottleneck module



(a) YOLOv5s-Ghost

(b) YOLOv5s-Ghost-SE

图 5 YOLOv5s-Ghost 与 YOLOv5s-Ghost-SE 网络结构

Fig. 5 The structures of the YOLOv5s-Ghost and YOLOv5s-Ghost-SE

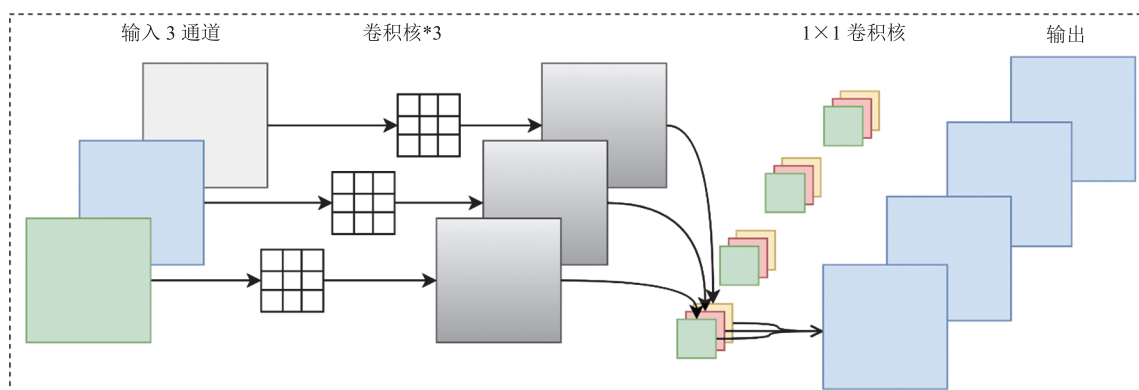


图 6 深度可分离卷积网络结构

Fig. 6 Structure of the depthwise separable CNN

表 4 不同特征提取网络目标检测算法性能

Table 4 The performance comparison with different feature extraction networks

	平均精度 mAP (%)	每秒处理帧数 FPS (帧/s)	精准率 (%)	召回率 (%)	每秒十亿次浮点运算数 GFLOPS (十亿次/s)	模型参数存储大小 (MB)
YOLOv5s	79.88	6	86.46	72.16	16.5	13.6
YOLOv5-0.40	79.53	8	86.18	71.50	11.1	9.16
YOLOv5s-MobileNetV3	78.14	4	86.02	70.03	7.1	6.7
YOLOv5s-Ghost	78.42	9	86.15	70.35	6.1	5.4
YOLOv5s-Ghost-SE	61.72	8	83.25	56.14	6.9	6.1

表 5 重构推理网络后的性能指标

Table 5 The performance comparison with different modified inference networks

	平均精度 mAP (%)	每秒处理帧数 FPS (帧/s)	精准率 (%)	召回率 (%)	每秒十亿次浮点运算数 GFLOPS (十亿次/s)	模型参数存储大小 (MB)
YOLOv5s-MobileNetV3	72.25	7	47.57	75.83	5.9	3.64
YOLOv5s-Ghost	78.16	12	85.50	70.12	3.4	2.37
YOLOv5n	75.70	9	85.16	67.76	4.2	3.65
YOLOv5s-Ghost-SE	57.26	6	81.22	52.19	4.5	2.94

检测精度仅小幅下降，且模型推理速度也有所提高。

4.3 重构推理网络

进一步对推理网络进行轻量化设计。与重构特征提取网络相同，使用深度可分离卷积替换原始推理网络中的普通卷积操作，使用 GhostBottleneck 模块替换原始推理网络中的 C3 模块，将重构后的推理网络结合上述 3 种不同特征提取网络进行实验分析，且不再进行重命名，

重构推理网络后的各项性能指标如表 5 所示。

由表 5 可知，重构推理网络后，与原始网络 YOLOv5s 相比，YOLOv5s-MobileNetV3 的 mAP 下降了 7.63%，推理速度仅提高了一帧；重构推理网络后的 YOLOv5s-Ghost-SE 虽然精准率较高，但 mAP 较原始网络 YOLOv5s 下降了 22.62%，推理速度较表 4 并未增加；YOLOv5s-Ghost 的检测精度相较于原始网络略微下降，但浮点数压缩为原来的 2.22 倍，模型大小压缩

为原来的 2.6 倍, 是 3 种设计方案中最小的。此外, YOLOv5s-Ghost 将相同 CPU 上的推理速度提升为原来的两倍, 又因为网络结构的轻量化, 在相同平台、数据集和训练参数下, YOLOv5s-Ghost 的训练时长小于原始网络, 各个性能指标均优于 YOLOv5n。经实验分析, 确定 YOLOv5s-Ghost 为最优重构方案。

YOLOv5s-Ghost 将 GhostBottleneck 模块与 DSConv 相结合, 重构 YOLOv5s 网络结构。使用 GhostBottleneck 模块替换原始网络中大部分的 C3 模块, 输入端卷积层保持不变, 将其余卷积层改为 DSConv, 大大降低了模型的浮点数计算量, 减小了模型的体积, 且保证了模型的检测性能。

4.4 算法性能优化

针对无人机采集的图像中目标占比较小、识别准确率差的问题, 本实验对网络的 SPP 结构进行改进。原始的 SPP 结构通过 3 个池化窗口进行池化, 然后在通道的维度上进行合并, 如图 7 所示。

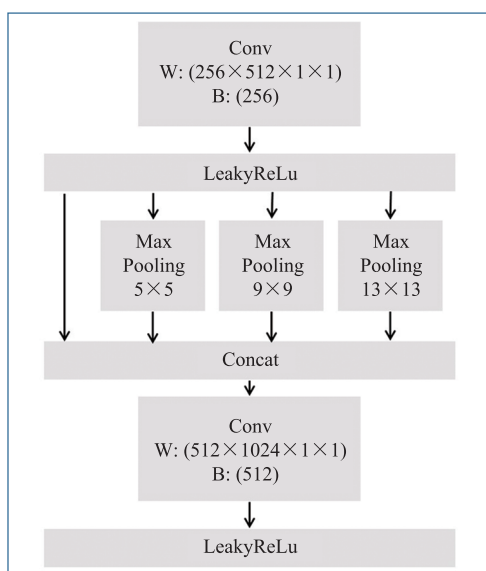


图 7 原始 SPP 结构

Fig. 7 The original structure of the SPP

为增加感受野, 使得网络对小目标的检测效果良好, 本研究增加 3×3 和 7×7 两个池化窗口, 改进后的网络结构如图 8 所示。为保证 SPP 模块的输出与 CSP 模块的兼容性, 本研究

将 SPP 模块中的第 2 个卷积层的权重矩阵参数由 $(512 \times 1024 \times 1 \times 1)$ 改为 $(512 \times 1536 \times 1 \times 1)$ 。

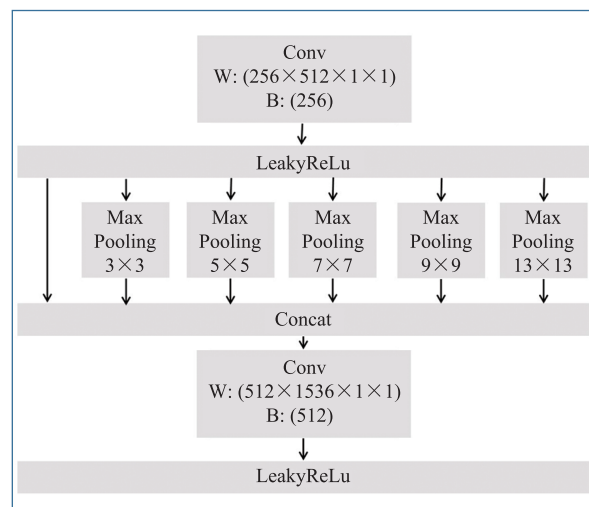


图 8 改进后的 SPP 结构

Fig. 8 The improved structure of the SPP

目前, 对于安全帽检测的研究大多针对网络结构进行改进, 对于推理时图像预处理的研究较少。由于光照对图像影响较大, 所以本研究在图像采集时对图像进行直方图均衡化处理, 以降低光照对图像的影响。

为提高目标检测模型对安全帽的检测准确性, 本文使用 CIoU 损失参数替换 GIoU 损失参数 (如公式 (5) 所示), 另外还使用 Focal loss 替换二进制交叉熵损失参数。Kaiming 等^[23]针对样本类别不均衡会导致准确率下降的问题, 提出了 Focal loss 损失函数 (公式 (7)), 该损失函数是根据标准交叉熵损失改进而来。通过减少易分类样本的权重, 使模型在训练时更专注于难分类样本。

$$L_{CIoU} = 1 - IOU(A, B) + \frac{Distance_2^2}{Distance_C^2} + \frac{v^2}{(1 - IOU(A, B) + v)} \quad (5)$$

其中, $Distance_2$ 为两个框中心点的欧氏距离; $Distance_C$ 为两个框最小外接矩形的对角线长度; v 为衡量长宽比一致性的参数, 其定义如公

式(6)所示。

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^p}{h^p} \right) \quad (6)$$

其中, w^{gt} 为真实框的宽; h^{gt} 为真实框的长; w^p 为预测框的宽; h^p 为预测框的长。

$$L_{if} = \begin{cases} -\alpha(1-y')^\gamma \log y', y=1 \\ -(1-\alpha)y'^\gamma \log(1-y'), y=0 \end{cases} \quad (7)$$

其中, y' 为网络输出的类别概率值; y 为真实标签; γ 为调节简单样本权重降低的速率, 当 γ 为 0 时, 公式(7)即为交叉熵损失函数, 当 γ 增加时, 调整因子的影响也增加; α 为平衡因子, 用以平衡正负样本本身的比例不均。

使用 DIOU_NMS 替换 YOLOv5 原有的 NMS, 使得网络在目标重叠时仍然可以具有较好的检测效果。

5 真实场景测试

自行搭建无人机平台, 使用 PX4 开源飞控, 搭载相机云台捕获施工工地图像信息并传输至计

算机进行处理。无人机可实现一键起飞、定点巡航、自动返航功能^[24-25]。无人机图片如图 9 所示, 具体硬件配置如表 6 所示。

表 6 无人机参数配置表

Table 6 Details of the applied UAV system

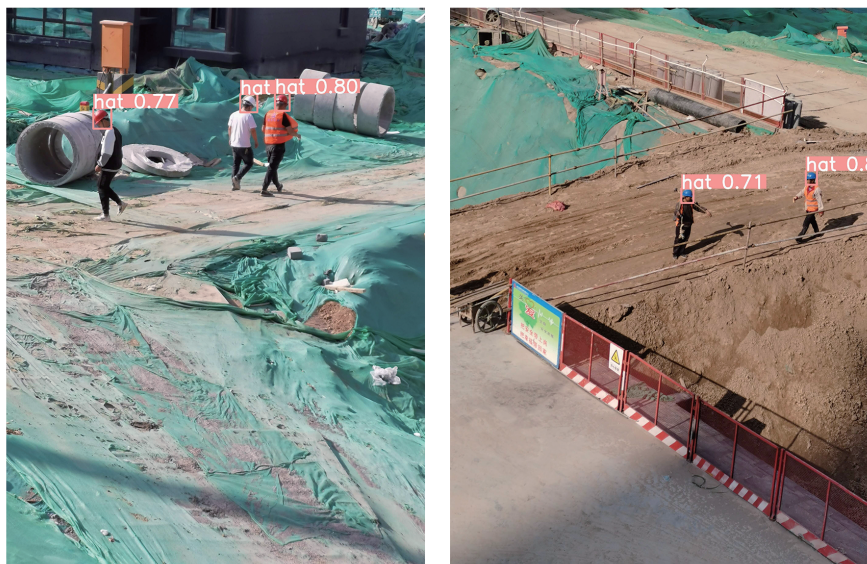
参数	配置
控制器	PixHawk 4
固件	PX4 1.11.0rc
地面站	QGroundControl V3.5.6
数据传输模块	Holybro Telemetry Radio 433MHz 100mW
GPS	M8N
电池	格氏 GS-00mAh-30C-4S1P
相机	飞萤 8SE

分别使用本研究设计的轻量化安全帽检测算法和原始 YOLOv5, 对无人机采集的图像信息进行检测, 检测结果如图 10 所示。由图 10 可知, 轻量化后的网络模型检测效果良好, 与原始网络基本保持同一水平, 但模型计算量和参数量远小于原始模型, 相同 CPU 检测速度较原始模型提高了 6 帧, 验证了本研究设计的轻量化智能目标检测的有效性和通用性。



图 9 无人机细节展示

Fig. 9 Details of the applied UAV



(a) YOLOv5 检测无人机拍摄图像



(b) YOLOv5s-Ghost 检测无人机拍摄图像

图 10 无人机实际量测图及采用不同方法获得的检测结果

Fig. 10 The results of practical measurements from a UAV using different detection method

6 结 论

本文结合目标检测算法与无人机技术, 提出了基于巡逻无人机的轻量型安全帽佩戴检测方法, 可对工人安全帽佩戴情况进行实时、自动、全方位的检测, 解决了固定摄像头和人工巡逻所带来的监控区域固定、成本高、效率低

的问题。通过实验验证, 该系统定点飞行、一键起飞、自动返航功能效果稳定, 图像与数据传输稳定, 改进后的算法检测精度评估指标——mAP 与 YOLOv5s 并无较大差别(仅下降 1.72%), 各项功能均达到预期效果, 并在当地多个建筑工地进行测试使用。但由于时间和精力限制, 该系统也存在一些不足, 如对轻量

化目标检测模型,可以直接使用树莓派或工控机等设备在无人机上进行处理,仅将检测结果传回地面站计算机。此外,该模型无法对检测到的未佩戴安全帽的员工进行跟踪飞行。后期工作计划将对相关不足进行优化,并进行推广应用。

参 考 文 献

- [1] 朱富云,戴相龙,傅靖,等. 基于视频监控的变电站内安全帽检测算法研究 [J]. 电网与清洁能源, 2018, 34(12): 71-76.
Zhu FY, Dai XL, Fu J, et al. Research on hard hat detection algorithm in substation based on video surveillance [J]. Power System and Clean Energy, 2018, 34(12): 71-76.
- [2] Waranusast R, Bundon N, Timtong V, et al. Machine vision techniques for motorcycle safety helmet detection [C] // Proceedings of the 2013 28th International Conference on Image and Vision Computing New Zealand, 2013: 35-40.
- [3] Rubaiyat AHM, Toma TT, Kalantari-Khandani M, et al. Automatic detection of helmet uses for construction safety [C] // Proceedings of the 2016 IEEE/WIC/ACM International Conference on Web Intelligence Workshops, 2016: 135-142.
- [4] 王宏任,陈世峰. 基于关键点检测二阶段目标检测方法研究 [J]. 集成技术, 2021, 10(5): 34-42.
Wang HR, Chen SF. Research on two-stage object detection method based on key point detection [J]. Journal of Integration Technology, 2021, 10(5): 34-42.
- [5] 吴庆甜,郭会文,吴新宇,等. 基于巡逻机器人的实时跑动检测系统 [J]. 集成技术, 2017, 6(3): 50-58.
Wu QT, Guo HW, Wu XY, et al. Real-Time running detection from a patrol robot [J]. Journal of Integration Technology, 2017, 6(3): 50-58.
- [6] Wang Y, Shan GB, Li H, et al. A wearable-sensor system with AI technology for real-time biomechanical feedback training in hammer throw [J]. Sensors, 2023, 23(1): 425.
- [7] Li P, Wang L, Zhu GL, et al. Predictive active disturbance rejection control for servo systems with communication delays via sliding mode approach [J]. IEEE Transactions on Industrial Electronics, 2021, 68(12): 12679-12688.
- [8] 张勇,吴孔平,高凯,等. 基于改进型 YOLOv3 安全帽检测方法的研究 [J]. 计算机仿真, 2021, 38(5): 413-417.
Zhang Y, Wu KP, Gao K, et al. Helmet detection based on modified YOLOv3 [J]. Computer Simulation, 2021, 38(5): 413-417.
- [9] 孙国栋,李超,张航. 融合自注意力机制的安全帽佩戴检测方法 [J]. 计算机工程与应用, 2022, 58(20): 300-304.
Sun GD, Li C, Zhang H. Safety helmet wearing detection method fused with self-attention mechanism [J]. Computer Engineering and Applications, 2022, 58(20): 300-304.
- [10] 张明媛,曹志颖,赵雪峰,等. 基于深度学习的建筑工人安全帽佩戴识别研究 [J]. 安全与环境学报, 2019, 19(2): 535-541.
Zhang MY, Cao ZY, Zhao XF, et al. Research on the recognition of construction workers' helmet wearing based on deep learning [J]. Journal of Safety and Environment, 2019, 19(2): 535-541.
- [11] 王忠玉. 智能视频监控下的安全帽佩戴检测系统的设计与实现 [D]. 北京:北京邮电大学, 2018.
Wang ZY. Design and implementation of detection safety of wearing helmet based on intelligent video surveillance [D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
- [12] 徐守坤,王雅如,顾玉宛,等. 基于改进 Faster RCNN 的安全帽佩戴检测研究 [J]. 计算机应用研究, 2020, 37(3): 901-905.
Xu SK, Wang YR, Gu YW, et al. Safety helmet wearing detection study based on improved Faster RCNN [J]. Application Research of Computer, 2020, 37(3): 901-905.
- [13] 夏鼎,王亚立,乔宇. 基于人体部件的视频行为识别方法研究 [J]. 集成技术, 2021, 10(5): 23-33.
Xia D, Wang YL, Qiao Y. Research of video action recognition based on human body parts [J]. Journal of Integration Technology, 2021, 10(5): 23-33.

- [14] 王雨生, 顾玉宛, 庄丽华, 等. 复杂姿态下的安全帽佩戴检测方法研究 [J]. 计算机工程与应用, 2022, 58(1): 190-196.
Wang YS, Gu YW, Zhuang LH, et al. Research on the detection method of helmet wearing in complex posture [J]. Computer Engineering and Applications, 2022, 58(1): 190-196.
- [15] 韩锬, 李斯宇, 肖友刚. 施工场景下基于 YOLOv3 的安全帽佩戴状态检测 [J]. 铁道科学与工程学报, 2021, 18(1): 268-276.
Han K, Li SY, Xiao YG. Detection of wearing state helmet based on YOLOv3 in construction scene [J]. Journal of Railway Science and Engineering, 2021, 18(1): 268-276.
- [16] 陈星辉, 赛景波, 李新梦, 等. 基于无人机航拍的油气管道监管系统设计与实现 [J]. 集成技术, 2019, 8(4): 52-62.
Chen XH, Sai JB, Li XM, et al. Design and implementation of oil and gas pipeline monitoring system based on aerial photography of unmanned aerial vehicle [J]. Journal of Integration Technology, 2019, 8(4): 52-62.
- [17] Ciaparrone G, Sánchez FL, Tabik S, et al. Deep learning in video multi-object tracking: a survey [J]. Neurocomputing, 2020, 381: 61-88.
- [18] Liu P, Chen AY, Huang YN, et al. A review of rotorcraft unmanned aerial vehicle (UAV) developments and applications in civil engineering [J]. Smart Structures and Systems, 2014, 13(6): 1065-1094.
- [19] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [20] Howard A, Sandler M, Chu G, et al. Searching for MobileNetV3 [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314-1324.
- [21] Han K, Wang YH, Tian Q, et al. GhostNet: more features from cheap operations [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.
- [22] Chollet F. Xception: deep learning with depthwise separable convolutions [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.
- [23] Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection [C] // Proceedings of the 2017 IEEE International Conference on Computer Vision, 2017: 2999-3007.
- [24] 王丹爽. 无人机技术在施工现场安全评价中的应用研究 [D]. 哈尔滨: 哈尔滨工业大学, 2017.
Wang DS. Application of UAV technology in construction site safety evaluation [D]. Harbin: Harbin Institute of Technology, 2017.
- [25] 李彩霞. 无人机在建筑施工现场的应用研究 [D]. 哈尔滨: 哈尔滨工业大学, 2016.
Li CX. Research on application of UAVs for construction site management [D]. Harbin: Harbin Institute of Technology, 2016.