

文章编号:1671-251X(2020)11-0054-05

DOI:10.13272/j.issn.1671-251x.2020050033

多尺度特征融合的煤矿救援 机器人目标检测模型

翟国栋¹, 任聪¹, 王帅², 岳中文³, 潘涛^{4,5,6}, 季如佳¹

(1. 中国矿业大学(北京)机电与信息工程学院, 北京 100083;

2. 湖北工业大学电气与电子工程学院, 湖北 武汉 430068;

3. 中国矿业大学(北京)力学与建筑工程学院, 北京 100083;

4. 神华信息技术有限公司, 北京 100011; 5. 智能矿山(煤炭行业)工程研究中心, 北京 100011;

6. 闽江学院福建省信息处理与智能控制重点实验室, 福建 福州 350121)



扫码移动阅读

摘要:传统目标检测模型采用人工设计的目标特征,造成检测精度较差。基于深度学习的目标检测模型具有较高的检测精度,然而针对实时性和精度要求比较高的煤矿救援机器人应用场合,获取的图像信息较少且目标特征不明显,造成目标检测效果较差。为提高目标检测精度和速度,基于YOLO V3模型提出了一种多尺度特征融合的煤矿救援机器人目标检测模型。该模型主要包括特征提取和特征融合2个模块:特征提取模块采用空洞瓶颈和多尺度卷积获得更加丰富的图像特征信息,增强目标特征表达能力,提高了目标分类精度和检测速度;特征融合模块在特征金字塔中引入空间注意力机制,对含有丰富语义信息的高层特征图和含有丰富位置信息的低层特征图进行有效融合,弥补了高层特征图位置信息表达能力不足的缺点,提高了目标定位精度。将该模型部署在煤矿救援机器人嵌入式NVIDIA Jetson TX2平台上进行灾后环境目标检测实验,检测精度为88.73%,检测速度为28帧/s,满足煤矿救援机器人目标检测的实时性和精度需求。

关键词:煤矿救援机器人; 目标检测; 多尺度特征融合; YOLO V3; 深度学习

中图分类号:TD67 文献标志码:A

Object detection model of coal mine rescue robot based on multi-scale feature fusion

ZHAI Guodong¹, REN Cong¹, WANG Shuai², YUE Zhongwen³, PAN Tao^{4,5,6}, JI Rujia¹

(1.School of Mechanical Electronic and Information Engineering, China University of Mining and Technology(Beijing), Beijing 100083, China; 2.School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, China; 3.School of Mechanics and Civil Engineering, China University of Mining and Technology(Beijing), Beijing 100083, China; 4.Shenhua Information Technology Co., Ltd., Beijing 100011, China; 5.Intelligent Mine (Coal Industry) Engineering Research Center, Beijing 100011, China; 6.Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, Fuzhou 350121, China)

Abstract:Traditional object detection model uses artificial object features, resulting in poor detection accuracy. Object detection model based on deep learning has high detection accuracy. However, for application of coal mine rescue robot with high real-time and accuracy requirements, the obtained image

收稿日期:2020-05-14;修回日期:2020-11-03;责任编辑:盛男。

基金项目:国家重点研发计划项目(2017YFC0804307);中央高校基本科研业务费专项资金资助项目(2020YJSJD06);福建省信息处理与智能控制重点实验室(闽江学院)开放基金项目(MJUKF-IPIC201905)。

作者简介:翟国栋(1973—),男,河北高碑店人,副教授,博士,主要研究方向为煤矿机械和除尘设备设计,E-mail:zgd@cumtb.edu.cn。通信作者:任聪(1996—),男,湖北襄阳人,硕士研究生,主要研究方向为图像处理与机器人同步定位与建图,E-mail:rencong_jx@163.com。

引用格式:翟国栋,任聪,王帅,等.多尺度特征融合的煤矿救援机器人目标检测模型[J].工矿自动化,2020,46(11):54-58.

ZHAI Guodong,REN Cong,WANG Shuai,et al. Object detection model of coal mine rescue robot based on multi-scale feature fusion [J]. Industry and Mine Automation,2020,46(11):54-58.

information is less and object features are not obvious, resulting in poor object detection effect. In order to improve accuracy and speed of object detection, on the basis of YOLO V3 model, an object detection model of coal mine rescue robot based on multi-scale feature fusion is proposed. The model mainly includes feature extraction module and feature fusion module. The feature extraction module uses hole bottleneck and multi-scale convolution to obtain more abundant image feature information, so as to enhance expression ability of object feature and improve object classification accuracy and detection speed. The feature fusion module introduces spatial attention mechanism into feature pyramid to effectively fuse high-level feature map with rich semantic information and low-level feature map with rich location information, which makes up for lack of position information expression ability of high-level feature map, and improves object positioning accuracy. The model is deployed on embedded NVIDIA Jetson TX2 platform in coal mine rescue robot for post-disaster environment object detection experiment. The detection accuracy is 88.73% and detection speed is 28 frames per second, which meet real-time and precision requirements of object detection of coal mine rescue robot.

Key words: coal mine rescue robot; object detection; multi-scale feature fusion; YOLO V3; deep learning

0 引言

煤矿灾害事故具有灾难性、突发性、继发性和破坏性等特点,救灾时效性强,但救援人员进入井下施救易遇到二次爆炸等危险状况^[1-3]。为保障救援人员安全和提高救援效率,可在发生煤矿灾害事故后第一时间安排救援机器人进入灾害现场,进行灾后复杂环境的信息采集和相关救援工作。

煤矿灾后环境复杂,救援机器人要想完全代替人员自主完成搜救任务,首先需要通过相机实时采集灾后环境图像,然后根据图像对周围环境进行解析,最后借助视觉检测算法进行目标检测,以实现自主运动。传统的目标检测模型^[4-5]首先通过不同大小的滑动窗口从图像中生成目标候选框,然后在目标候选框内进行特征提取,最后对提取的特征进行分类。该模型采用基于滑动窗口的候选框生成策略,时间复杂度高,检测速度较慢,且提取的基本是人工设计的目标特征,对于目标变化没有很好的鲁棒性,导致检测精度较差。基于深度学习的目标检测模型能够自适应地从大量图像中学习目标特征,具有检测精度高和鲁棒性好等优点,已逐步在不同领域的目标检测任务中广泛应用。基于深度学习的目标检测模型主要分为两阶段目标检测模型^[6-8]和单阶段目标检测模型^[9]。两阶段目标检测模型首先生成 1 组稀疏的目标候选框,然后通过卷积神经网络进行目标分类和位置回归,最后输出目标的检测类别和位置坐标;该类模型具有较高的检测精度,但在目标候选框生成上需要消耗大量时间,不适合部署在具有实时检测需求的应用场景中。单阶段目标检测模型将目标检测问题转换为回归问题进行处理,不需要生成目标候选框,具有检测速度快的优

点,但在煤矿救援机器人应用场景下,获取的图像信息较少且目标特征不明显,导致目标检测效果较差。

本文在主流的单阶段目标检测模型 YOLO V3 的基础上,提出了一种多尺度特征融合的煤矿救援机器人目标检测模型。该模型利用特征提取模块获得输入图像特征,得到不同尺度的特征图,再通过特征融合模块实现不同尺度特征图的有效融合,从而提高目标检测精度和速度。

1 多尺度特征融合的煤矿救援机器人目标检测模型

1.1 特征提取模块

基于深度分组卷积提出了一种轻量型特征提取模块,用来替代 YOLO V3 中的 Darknet-53^[10],结构见表 1,其中 c 为特征图通道参数,用来平衡模型检测精度和速度。特征提取模块由 5 个阶段连接起来构成,其中阶段 2—阶段 5 均由 1 个空洞瓶颈子模块、多个多尺度卷积子模块和 1 个下采样子模块组成。

(1) 初始化子模块。在阶段 1 中,受 Inception V4^[11] 和 DSOD (Deeply Supervised Object Detector)^[12] 的启发,设计初始化子模块,用于对输入图像进行 4 倍下采样并扩展特征图通道数至 32,如图 1 所示。首先将输入图像经过步长为 2 的 3×3 卷积层进行 2 倍下采样;然后采用双通道来获得不同大小的感受野,一个通道采用 1×1 卷积层和 3×3 卷积层堆叠来学习大尺度目标的特征,另一个通道采用步长为 2 的最大池化层来学习小尺度目标的特征;最后通过通道合并和 1×1 卷积层控制模型维度,减少模型计算量。

(2) 空洞瓶颈子模块。为获得丰富的特征信息,在空洞卷积^[13]的基础上构建空洞瓶颈子模块,如图 2 所示。首先上一层特征图通过 1×1 卷积层

表 1 特征提取模块结构
Table 1 Structure of feature extraction module

阶段	类型	参数	输出尺寸
阶段 1	初始化子模块	3×3 卷积层/池化层, 步长为 2	$104\times 104\times 32$
阶段 2	空洞瓶颈子模块	—	$104\times 104\times 4c$
	多尺度卷积子模块	多尺度卷积子模块 $\times 3$	$104\times 104\times 4c$
	下采样子模块	1×1 卷积层, 步长为 1	$104\times 104\times 4c$
		2×2 平均池化层, 步长为 2	$52\times 52\times 4c$
阶段 3	空洞瓶颈子模块	—	$52\times 52\times 8c$
	多尺度卷积子模块	多尺度卷积子模块 $\times 4$	$52\times 52\times 8c$
	下采样子模块	1×1 卷积层, 步长为 1	$52\times 52\times 8c$
		2×2 平均池化层, 步长为 2	$26\times 26\times 8c$
阶段 4	空洞瓶颈子模块	—	$26\times 26\times 10c$
	多尺度卷积子模块	多尺度卷积子模块 $\times 6$	$26\times 26\times 10c$
	下采样子模块	1×1 卷积层, 步长为 1	$26\times 26\times 10c$
		2×2 平均池化层, 步长为 2	$13\times 13\times 10c$
阶段 5	空洞瓶颈子模块	—	$13\times 13\times 12c$
	多尺度卷积子模块	多尺度卷积子模块 $\times 3$	$13\times 13\times 12c$
	下采样子模块	1×1 卷积层, 步长为 1	$13\times 13\times 12c$
		2×2 平均池化层, 步长为 1	$13\times 13\times 12c$

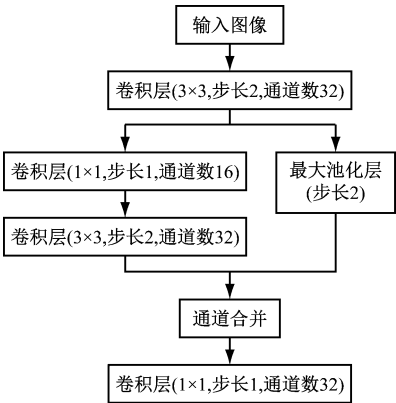


图 1 初始化子模块结构

Fig. 1 Structure of stem block submodule

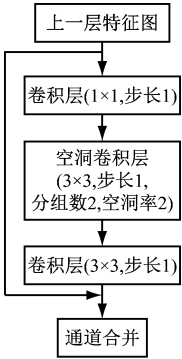


图 2 空洞瓶颈子模块结构

Fig. 2 Structure of hole bottleneck submodule

进行特征图通道压缩;其次采用分组数为 2 的 3×3 空洞卷积层进行特征提取,以提高模型计算效率和增大浅层网络的感受野,保留目标的特征信息;然后

利用 3×3 卷积层增强目标特征语义信息;最后利用残差连接将层次较低的上一层特征图与经过 3×3 卷积层后的特征图相结合,增强目标特征表达能力。

(3) 多尺度卷积子模块。在 ResNet^[14] 结构的基础上,利用多尺度卷积增强网络特征提取能力,构建多尺度卷积子模块,如图 3 所示。首先通过 1×1 卷积层对上一层特征图通道进行压缩;其次均匀地将特征图分成 n 个特征图子集 $X_1, X_2, \dots, X_n$, 其中每个特征图子集与输入特征图具有相同的大小,但

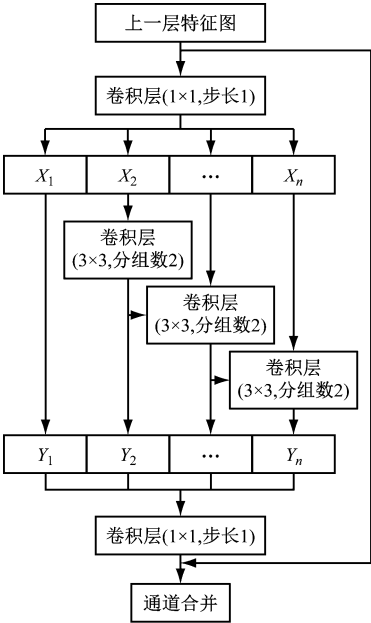


图 3 多尺度卷积子模块结构

Fig. 3 Structure of multi-scale convolution submodule

特征图通道数为 $1/n$;然后在单个特征图子集内利用分组数为 2 的 3×3 卷积层进行特征增强,实现同一特征图子集内不同通道的信息融合,以提高目标分类精度;接着,在不同特征图子集间采用特征金字塔融合方式,即将特征图子集与上一个特征图子集求和后输入到分组数为 2 的 3×3 卷积层中进行运算,实现不同特征图子集间的信息融合;最后利用 1×1 卷积层对 n 个特征图子集输出的特征信息 $Y_1, Y_2, \dots, Y_n$ 进行充分融合来增强特征表达能力。

1.2 特征融合模块

在 YOLO V3 模型特征金字塔的基础上嵌入空间注意力机制^[15],构建特征融合模块,如图 4 所示。将高层特征图通过 1×1 卷积层实现特征图通道压缩;利用步长为 2 的 2×2 反卷积层对特征图进行上采样,上采样后的特征图通过空间注意力机制直接对位置特征进行增强,提高目标定位精度;将含有丰富位置信息的低层特征图通过 3×3 卷积层进行特征提取后与含有丰富语义信息的高层特征图进行通道合并,用于后续的目标分类与目标定位。

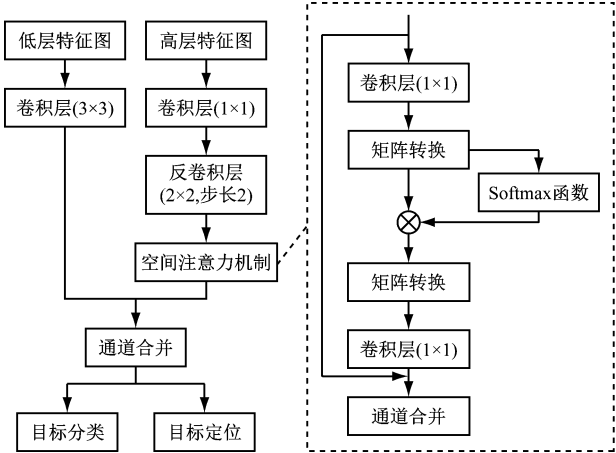


图 4 特征融合模块结构

Fig. 4 Structure of feature fusion module

2 实验及结果分析

利用 Python/C++ 联合开发多尺度特征融合的煤矿救援机器人目标检测模型,基于 Ubuntu18.04 搭建模型训练平台:CPU 为 I5-9600K, RAM 为 32 GB, GPU 为 NVIDIA GTX 1660 SUPER。设置模型检测阈值为 0.5,特征提取模块中特征图通道参数 c 为 32,特征图子集数 n 为 4。采用迁移学习进行模型训练,主要分为预训练和迁移微调:预训练采用 COCO 数据集和 VOC 数据集中行人目标进行训练,得到模型的初始权重;迁移微调利用预训练阶段产生的初始权重作为输入,在煤矿行人检测数据集上进行微调训练,使模型能够充分学习煤矿环境下的行人目标特征。采用随机梯度下降法^[16]和反向传播^[17]来优化模型,权重衰减为 0.005,动量为 0.9,总迭代次数为 20 000 次,初始学习率为 0.01,

当迭代至 18 000 次时,学习率减小至 0.001 使模型收敛。

在煤矿救援机器人嵌入式 Jetson TX2 平台上分别部署 YOLO V3 模型和多尺度特征融合的煤矿救援机器人目标检测模型进行测试,结果见表 2。可看出与 YOLO V3 模型相比,多尺度特征融合的煤矿救援机器人目标检测模型的检测精度和速度均有所提高。

表 2 目标检测模型测试结果对比

Table 2 Comparison of test results of object detection models

目标检测模型	检测精度/%	检测速度/(帧·s ⁻¹)
YOLO V3 模型	79.43	11
多尺度特征融合的煤矿救援机器人目标检测模型	88.73	28

针对煤矿灾后昏暗、灰尘和烟雾环境,YOLO V3 模型和多尺度特征融合的煤矿救援机器人目标检测模型检测结果如图 5 所示。可看出 YOLO V3 模型可检测出部分行人,但在行人密度较大时存在大量漏检和误检现象;多尺度特征融合的煤矿救援机器人目标检测模型在不同环境下均可准确预测行人位置,具有更好的鲁棒性。

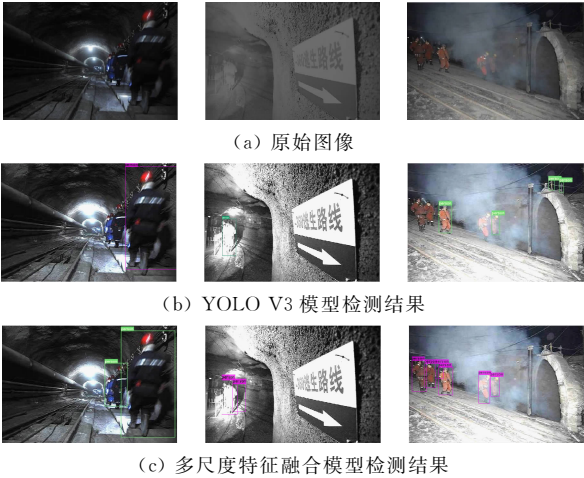


图 5 煤矿灾后环境目标检测结果对比

Fig. 5 Comparison of object detection results in coal mine post-disaster environment

3 结语

基于 YOLO V3 模型提出了一种多尺度特征融合的煤矿救援机器人目标检测模型。该模型主要包括特征提取和特征融合 2 个模块:特征提取模块采用空洞瓶颈和多尺度卷积生成含有丰富信息的不同尺度的特征图,增强目标特征表达能力,提高了目标分类精度和检测速度;特征融合模块在特征金字塔中引入空间注意力机制,对含有丰富语义信息的高层特征图和含有丰富位置信息的低层特征图的信息

进行充分融合,弥补了高层特征图位置信息表达能力不足的缺点,有效提高了目标定位精度。将该模型部署在煤矿救援机器人嵌入式 NVIDIA Jetson TX2 平台上进行灾后环境目标检测实验,结果表明该模型具有较高的目标检测精度和速度,鲁棒性好,满足煤矿救援机器人目标检测的实时性和精度需求。

参考文献(References):

- [1] 李现国,李斌,刘宗鹏,等.井下视频行人检测方法[J].工矿自动化,2020,46(2):54-58.
LI Xianguo, LI Bin, LIU Zongpeng, et al. Underground video pedestrian detection method[J]. Industry and Mine Automation, 2020, 46(2): 54-58.
- [2] 郑学召,赵炬,张铎,等.煤矿救援机器人研究现状及发展趋势[J].工矿自动化,2019,45(9):7-12.
ZHENG Xuezhao, ZHAO Ju, ZHANG Duo, et al. Research status and development trend of coal mine rescue robot[J]. Industry and Mine Automation, 2019, 45(9): 7-12.
- [3] 葛世荣,朱华.危险环境下救援机器人技术发展现状与趋势[J].煤炭科学技术,2017,45(5):1-8.
GE Shirong, ZHU Hua. Technical development status and tendency of rescue robot in dangerous environment[J]. Coal Science and Technology, 2017, 45(5): 1-8.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, 2005: 886-893.
- [5] 高凯亮,覃团发,陈跃波,等.一种混合高斯背景模型下的像素分类运动目标检测方法[J].南京大学学报(自然科学版),2011,47(2):195-200.
GAO Kailiang, QIN Tuanfa, CHEN Yuebo, et al. Detection of moving objects using pixel classification based on gaussian mixture model[J]. Journal of Nanjing University(Natural Sciences), 2011, 47(2): 195-200.
- [6] 郭毓,苏鹏飞,吴益飞,等.基于Faster R-CNN的机器人目标检测及空间定位[J].华中科技大学学报(自然科学版),2018,46(12):55-59.
GUO Yu, SU Pengfei, WU Yifei, et al. Object detection and location of robot based on Faster R-CNN[J]. Journal of Huazhong University of Science and Technology(Natural Science Edition), 2018, 46(12): 55-59.
- [7] 彭秋辰,宋亦旭.基于Mask R-CNN的物体识别和定位[J].清华大学学报(自然科学版),2019,59(2):135-141.
PENG Qiuchen, SONG Yixu. Object recognition and localization based on Mask R-CNN[J]. Journal of Tsinghua University(Science and Technology), 2019, 59(2): 135-141.
- [8] 胡辉,曾琛.基于改进R-FCN的车辆目标检测方法[J].计算机工程与设计,2020,41(4):1164-1168.
HU Hui, ZENG Chen. Vehicle detection method based on improved R-FCN[J]. Computer Engineering and Design, 2020, 41(4): 1164-1168.
- [9] WANG Hongwei, LI Dahua, SONG Yu, et al. Single-shot object detection with split and combine blocks[J]. Applied Sciences, 2020, 10(18): 6382-6398.
- [10] XU Danqing, WU Yiquan. Improved YOLO-V3 with DenseNet for multi-scale remote sensing target detection[J]. Sensors, 2020, 20(15): 4276-4299.
- [11] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[C]//The 31st AAAI Conference on Artificial Intelligence, San Francisco, 2017: 4278-4284.
- [12] SHEN Zhiqiang, LIU Zhuang, LI Jianguo, et al. DSOD: learning deeply supervised object detectors from scratch[C]//IEEE International Conference on Computer Vision, Venice, 2017: 1937-1945.
- [13] 谢博,申国伟,郭春,等.基于残差空洞卷积神经网络的网络安全实体识别方法[J].网络与信息安全学报,2020,6(5):126-138.
XIE Bo, SHEN Guowei, GUO Chun, et al. Cyber security entity recognition method based on residual dilation convolution neural network[J]. Chinese Journal of Network and Information Security, 2020, 6(5): 126-138.
- [14] LI Baoqi, HE Yuyao. An improved ResNet based on the adjustable shortcut connections[J]. IEEE Access, 2018, 6: 18967-18974.
- [15] 雷俊锋,贺睿,肖进胜.融合空间注意力机制的行车障碍预测网络[J].光学精密工程,2020,28(8):1850-1860.
LEI Junfeng, HE Rui, XIAO Jinsheng. Driving obstacles prediction network merged with spatial attention[J]. Optics and Precision Engineering, 2020, 28(8): 1850-1860.
- [16] 王功鹏,段萌,牛常勇.基于卷积神经网络的随机梯度下降算法[J].计算机工程与设计,2018,39(2):441-445.
WANG Gongpeng, DUAN Meng, NIU Changyong. Stochastic gradient descent algorithm based on convolution neural network[J]. Computer Engineering and Design, 2018, 39(2): 441-445.
- [17] 徐凤强,董鹏,王辉兵,等.基于水下机器人的海产品智能检测与自主抓取系统[J].北京航空航天大学学报,2019,45(12):2393-2402.
XU Fengqiang, DONG Peng, WANG Huibing, et al. Intelligent detection and autonomous capture system of seafood based on underwater robot[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(12): 2393-2402.