

中文引用格式:纪执安,周云奕,张玉媛,等. 基于改进YOLOv5的工业现场不安全行为检测[J]. 中国安全科学学报,2024,34(7):38-43.

英文引用格式:JI Zhi'an,ZHOU Yunyi,ZHANG Yuyuan,et al. Industrial site unsafe behavior detection based on improved YOLOv5 [J]. China Safety Science Journal,2024,34(7):38-43.

基于改进YOLOv5的工业现场不安全行为检测*

纪执安 工程师,周云奕 工程师,张玉媛 工程师,郭新然 工程师,石坤 研究员
(中国特种设备检测研究院,北京100029)

中图分类号:X915.2 文献标志码:A DOI: 10.16265/j.cnki.issn1003-3033.2024.07.2030

基金项目:国家市场监督管理总局科技计划项目(2023MK208);中国特种设备检测研究院青年基金资助(2023youth07)。

【摘要】 为准确识别复杂工业现场人员的不安全行为,减少安全事故的发生,提出一种改进的YOLOv5不安全行为检测模型。首先,在YOLOv5的backbone部分引入注意力机制,提高卷积网络对不安全行为特征的敏感度;其次,通过图像几何变换和像素级处理丰富训练样本数量,提升检测模型在不同工业环境中的泛化能力;然后,蒸馏检测模型并优化网络结构参数来加速模型的训练;最后,将模型训练迭代200次,模拟起重吊索、机器人自动化产线和操作间3类工业现场,检测人员是否穿戴安全帽、工作服以及是否在安全区域工作,并依据行为划定危险等级,判定人员是否安全生产。结果表明:该模型能检测昏暗、光照和遮挡等多类复杂工业环境下人员的12种不安全行为,且在不安全行为测试集上精确率为98.6%,召回率为99.2%,平均精度为97.58%。

【关键词】 YOLOv5; 工业现场; 不安全行为; 检测模型; 注意力机制

Industrial site unsafe behavior detection based on improved YOLOv5

JI Zhi'an, ZHOU Yunyi, ZHANG Yuyuan, GUO Xinran, SHI Kun
(China Special Equipment Inspection and Research Institute, Beijing 100029, China)

Abstract: In order to accurately identify unsafe behaviors of personnel in complex industrial sites and reduce the occurrence of safety accidents, an improved YOLOv5 unsafe behavior detection model was proposed. Firstly, an attention mechanism was introduced in the backbone of YOLOv5 to enhance the sensitivity of convolutional networks to unsafe behavior features. Secondly, enriching the number of training samples through image geometric transformation and pixel-level processing enhanced the generalization ability of the detection model in different industrial environments. Then, the detection model was distilled, and the network structure parameters were optimized to accelerate the training of the mode. Finally, the model was trained and iterated 200 times to simulate three types of industrial sites: lifting slings, robot-automated production lines, and operating rooms. It detected whether personnel were wearing safety helmets, work clothes and working in safe areas, and determined the level of danger based on their behavior to ascertain whether they were working safely. The results show that the model can detect 12 types of unsafe behaviors of personnel in complex industrial environments, such as dim light, lighting, and occlusion. The accuracy on the unsafe behavior test set is 98.6%, the recall rate is 99.2%, and the average accuracy is 97.58%.

Keywords: YOLOv5; industrial site; unsafe behavior; detection model; attention module

0 引言

随着科学技术的迅猛发展,传统手工制造行业已逐渐被现代高精尖自动化产业所替代^[1],这使得工业现场存在较多机器人、快速传送带和大重型起重机等^[2],工作人员若不规范穿戴安全帽、工作服,或者非法闯入机器运行区域,可能会引发重大事故,危及个人生命安全。因此,检测复杂工业现场下人员的不安全行为显得十分必要^[3-4]。

行为检测是计算机视觉领域研究的热门课题之一^[5],早在 1977 年,计算机视觉理论就已被提出。经过几十年发展,计算机视觉技术广泛应用于视频监控、自动驾驶和工业制造等领域^[6-7]。近几年,机器学习的加入,让行为检测更具智能化^[8-9]。DALAL 等^[10]提出方向梯度直方图(Histogram Of Oriented Gradient, HOG)特征提取方法,通过支持向量机(Support Vector Machines, SVM)分类器检测人员行为。HUNG 等^[11]将迁移学习方法应用于 VGG19、Inception_V3 和 InceptionResnet_V2 等 3 个预训练模型,以此识别工业现场中的人员不安全行为。PATWAL 等^[12]使用轻量级卷积网络模型检测人员行为,并在化工安全生产上得到应用。2016 年,JOSEPH 等^[13]基于直接回归思路,提出 YOLO

(You Only Look Once)算法,该算法采用端到端的思想,将检测任务转化为单一的回归问题,提升了检测速度,被广泛应用于工业现场的不安全行为检测。

综上,现有研究在不安全行为检测方面多针对单类不安全行为,且不适应复杂工业现场的不安全行为检测。鉴于此,笔者拟改进 YOLOv5 算法,在原算法网络结构上融入双通道注意力机制和图像几何变化概念,使模型更聚焦于不安全行为关键特征,提高模型的特征表达能力。并基于改进后的模型,检测机器人自动化产线、起重吊索和操作间 3 类工业现场下的不安全行为,以期为复杂工业现场的事故预防提供新思路。

1 不安全行为数据采集

为使图像具有稳定性,减少噪声,在采集前固定相机,基于不同场景,改变不同拍摄角度,后经过图像帧间处理,确定不安全行为数据集为 5 000 张。主要涉及的工业场景包括起重吊索、机器人自动化产线、操作间等。并根据特定场景及人员行为类型,如是否戴安全帽、穿工作服,以及是否在安全区域,划定不同危险等级,设定 12 类不安全行为标签,见表 1。表 1 中,系数越高表示越危险。

表 1 不安全行为标签分类

Table 1 Unsafe behavior label classification

应用场景	不安全行为	危险系数	不安全行为标签
起重吊索	戴安全帽、穿工作服	安全	helmet, smock, safe
	戴安全帽、无工作服	危险 I	helmet, no_smock, dangerousI
	无安全帽、穿工作服	危险 I	no_helmet, smock, dangerousI
	无安全帽、无工作服	危险 II	no_helmet, no_smock, dangerousII
机器人自动化产线	安全区域、穿工作服	安全	safe_area, smock, safe
	安全区域、无工作服	危险 I	safe_area, no_smock, dangerousI
	危险区域、穿工作服	危险 I	dangerous_area, smock, dangerousI
	危险区域、无工作服	危险 II	dangerous_area, no_smock, dangerousII
工厂操作间	安全区域、工作服	安全	safe_area, smock, safe
	安全区域、无工作服	危险 I	safe_area, no_smock, dangerousI
	危险区域、穿工作服	危险 I	dangerous_area, smock, dangerousI
	危险区域、无工作服	危险 II	dangerous_area, no_smock, dangerousII

2 YOLOv5 模型结构

在 YOLOv4 结构上改进 YOLOv5 模型。相较于 YOLOv4,新增 C3 模块,C3 由 2 个卷积和 1 个残差网络构成,此模块能减少机器学习训练中的梯度消散问题,提高模型特征表达能力。从结构上来说,YOLOv5 由 3 个主干网络组成,即 backbone、neck 和

head,如图 1 所示。backbone 网络负责目标的初步特征提取。neck 网络通过上下采样把浅层的图像特征和更深层的语义特征相结合,并传递给预测网络。head 网络主要利用二元交叉熵损失函数回归预测输入图像。作为单阶段目标检测算法,YOLOv5 不需要像双阶段检测算法,事先提取图像中的目标作为候选区域,再从候选区域中进行分

类识别,而是直接生成检测目标的坐标和类别信息,这样大大提高了检测速度,更有利于在工程上的应用。

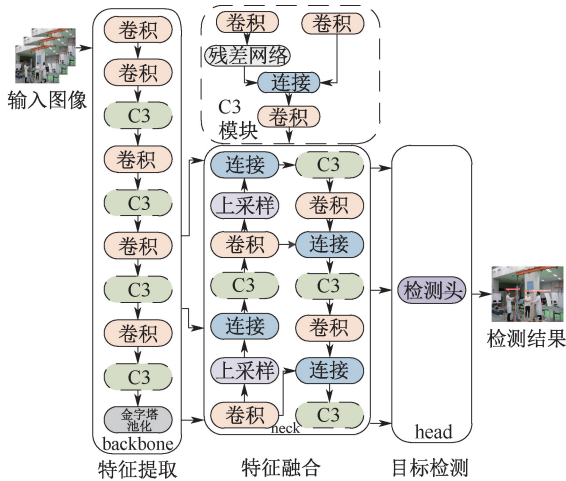


图1 YOLOv5 模型改进

Fig.1 YOLOv5 model structure

整个 YOLOv5 模型共有 5 个版本,包括 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 和 YOLOv5n。每个版本基础构造一致,差别在于参数 depth_multiple 和 width_multiple 设置不一,各版本适用于不同场景的速度和精度要求。为将模型方便部署在服务器上,选用深度较浅、检测速度较快的 YOLOv5s 作为基础模型。

3 改进 YOLOv5 模型

3.1 引入注意力机制

注意力机制^[14]与人类的选择性行为相似,其主要思想是在丰富的目标特征中选择更为重要的关键信息。在机器学习中,注意力机制主要应用于图像处理、数字和文本信息。在 backbone 中运用通道注意力和空间注意力相结合的方式改进模型结构。通道注意力模块关注于不安全行为特征图的通道维度,即不同通道上特征的重要性。通过计算每个通道的权重,加权特征图,从而增强对重要通道特征的关注,提高不安全行为特征的表达性。

空间注意力模块则关注于不安全行为特征图的空间维度,即不同位置上的特征重要性。通过计算特征图中每个位置的权重,突出重点空间区域,从而帮助模型关注到图像中最为重要的部分。空间注意力模块有助于模型捕捉不安全行为图像中的细节和关键信息,提高模型对空间信息的敏感性。

2 种注意力模块相结合,有利于加强各卷积层的数据信息联系,平衡各模块检测数据转换,减小内存消耗和信息过载问题,使模型更聚焦于不安全行为目标的关键特征,提高检测模型的特征表达能力,引入卷积注意力模块(Convolutional Block Attention Module, CBAM)的 YOLOv5 改进模型,如图 2 所示。

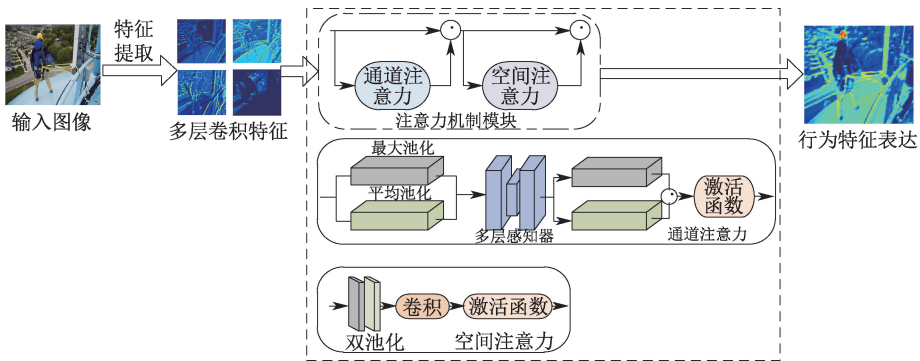


图2 引入 CBAM 的 YOLOv5 改进模型

Fig.2 Improved YOLOv5 model of incorporating CBAM

3.2 不安全行为数据集扩充

样本数量对于模型训练起着很大的作用,但在实际应用中,由于外界条件,不能采集各种复杂工业场景中的人员行为,因此,通过图像几何变换或像素级处理进行训练集扩充。如旋转、拼接、灰度化、增加白噪声等,将原有采集图片扩充到 12 000 张,以此来增加不安全行为的训练样本数量^[15],训练集扩充如图 3 所示。同时,为适应 YOLOv5 的模型训练,将采集样

本调整到 640×640 大小的图片,并通过 Labelimage 软件标注图像数据。划分机器人自动化产线、起重吊索和操作间 3 类不同场景,分别进行模型训练。

3.3 检测模型参数优化

YOLOv5 模型数据量大,结构复杂,卷积层较多。这样的结构是模型针对不同检测环境要求所设计。文中只针对自动化产线、起重吊索和操作间 3 类不同场景工业场景,因此,对模型蒸馏,可加速

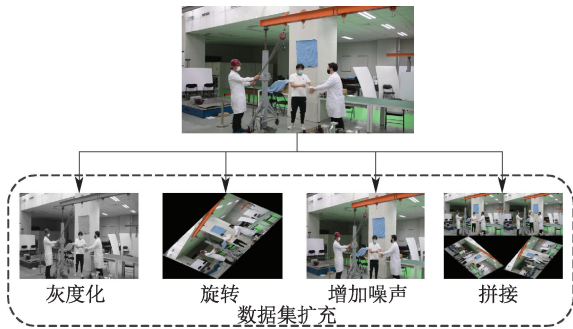


图3 训练集扩充

Fig. 3 Training set extension

模型的训练,同时也能提高目标检测速度。为此,引入贡献因子 γ ^[16] 来调节 neck 网络中的归一化层 (Batch Normalization, BN), 裁剪此模块中卷积贡献较小的结构。

$$\hat{Z} = \frac{Z_1 - \mu_\beta}{\sqrt{\sigma_\beta^2 + \varepsilon}} \quad (1)$$

$$Z_0 = \gamma \hat{Z} + \beta \quad (2)$$

式中: \hat{Z} 为 BN 层的过程预测值; Z_1 和 Z_0 为 BN 层的输入、输出信息; ε 为 BN 层的正则化项; μ_β 和 σ_β 为当前层 mini-batch 的平均值和方差; β 为缩放和偏移的仿射变换参数; γ 为 BN 层的贡献因子。在进行数据归一化时,当 γ 很小时,表示当前通道对模型训练影响小,可以裁掉。

4 不安全行为检测模型评价

不安全行为检测模型在中央处理器为 Intel Core i7-5960X, 图像处理器为 NVIDIA GeForce GTX 1080 Ti, 运行内存为 32 GB 随机存取存储器 (Random Access Memory, RAM) 的计算机上进行训练。设置训练批量大小为 16, 训练次数为 200 次, 学习率设置为 0.1%。

4.1 检测模型评价指标

检测模型的性能评价指标如下: 精确率 P 、召回率 R 和平均精度均值 (mean Average Precision, mAP)。 P 表示模型只识别相关目标的能力, 准确检测到所有数据结果的百分比, 其表示如下:

$$P = \frac{TP}{TP + FP} \quad (3)$$

式中: TP 为精确检测到的正样本数量; FP 为错误检测到的负样本数量。

R 表示在积极输入训练样本的情况下作出最理想预测的准确性, 指模型能够很好地检测到目标, 定

义如下式:

$$R = \frac{TP}{TP + FN} \quad (4)$$

式中 FN 为未检测到的正样本数量。

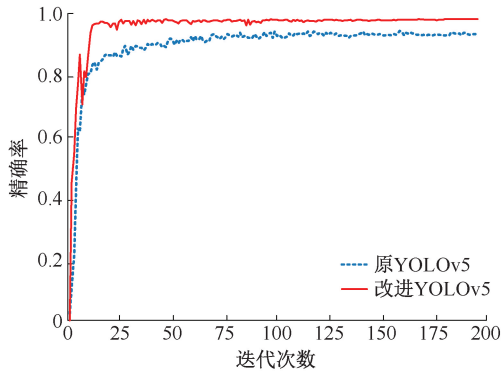
mAP 表示检测模型在多分类任务中的整体性能。在计算 mAP 时, 需独立计算每种不安全行为类别的平均精度 (Average Precision, AP), 后对所有不安全行为分类的 AP 值进行汇总并求取平均值。mAP@0.5 表示交并比 (Intersection over Union, IoU) 相除等于 0.5 时的平均精度。下式适用于对所有大场景的检测效果评价。

$$mAP = \frac{1}{N} \sum_{i=1}^N P(i) \quad (5)$$

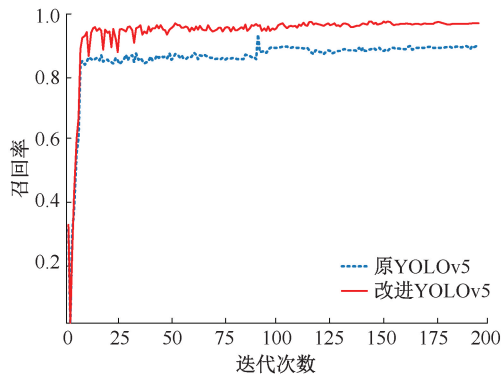
式中: $P(i)$ 为每种不安全行为检测的精度; N 为不安全行为类别的总数。

模型性能指标对比如图 4 所示。由图 4a 可知: 随着训练次数的增加, 原模型和改进模型在迭代 22 次时, 精确率均是迅速收敛上升, 在训练 100 次时, 2 模型精确率变化趋于平衡, 此时, 改进后的模型精确率为 98.6%, 原模型为 96.3%, 即改进后模型的精确率提升 2.4%。由图 4b 可知: 原模型和改进模型在训练迭代 130 次时, 召回率变化趋缓, 此时, 改进 YOLOv5 模型的召回率为 99.2%, 原模型的召回率为 95.9%, 改进后的 YOLOv5 模型较原模型, 召回率提升 3.4%。综合模型检测的精确率和召回率, 改进的 YOLOv5 模型整体性能比原模型有所提升。

为进一步分析不安全行为检测模型的性能, 选取现有经典行为识别算法, 快速区域卷积网络 (Fast Region-Convolutional Neural Network, Faster R-CNN)、YOLOv4、单步多框检测器 (Single Shot MultiBox Detector, SSD)、递归姿态注意力网络 (Recurrent Pose-Attention Network, RPAN)、时段网络 (Temporal Segment Networks, TSN) 和 YOLOv5s 进行对比, 不同算法平均精度见表 2, 评价指标为 mAP。在 3 类工业场景下, 影响目标检测的因素众多, 如起重机械、六轴自由度机械臂、栅栏、安全线和复杂背景, 因此, 对算法检测要求极高。在众多算法中, 改进后的 YOLOv5 模型的平均精度为 97.58%, 相较于现有 Faster R-CNN 算法提高 5.8%, 相比于 YOLOv4 算法提高 3.2%, 相较于 SSD 算法提高 7.1%, 相比于 RPAN 提高 4.9%, 与 TSN 相比提高 4.8%, 相比于 YOLOv5s 原始算法提高 2.3%, 在众多算法中, 改进后的 YOLOv5 算法识别效果最好。



(a) 精确率变化曲线



(b) 召回率变化曲线

图4 模型性能指标对比

Fig. 4 Comparison of model performance indicators

表2 不同算法平均精度

Table 2 Average accuracy of different algorithms

现有识别模型算法	mAP@ 0.5%
Faster R-CNN	92.21
YOLOv4	94.55
SSD	91.17
RPAN	92.97
TSN	93.15
YOLOv5s	95.36
改进后的 YOLOv5	97.58

4.2 工业场景应用

模拟起重吊索、机器人自动化产线和工厂操作间3类工业场景,不同工业环境下的不安全行为检测如图5所示。针对人员穿工作服、戴安全帽和所属工作区域,划分危险等级程度。在图5a左图中,涉及光照和背景干扰,工作人员正在借助吊索转移某重物,人员应该戴安全帽和穿工作服,此场景下,仅有1位作业人员穿戴符合规范,模型检测结果为 helmet, smock, safe; 有1人无安全帽,只穿戴工作服,危险程度等级为 I, 检测结果为 no_helmet, smock, dangerousI; 还有1位人员没有做任何安全防



图5 不同工业环境下的不安全行为检测

Fig. 5 Unsafe behavior detection in different industrial environments

护,危险程等级为 II,检测为 no_helmet, no_smock, dangerousII。图 5a 右图中也仅有 1 人符合安全操作规范,各目标检测置信度均在 95% 以上。在图 5b 机器人自动化产线场景下,涉及昏暗和遮挡干扰。图 5a 中,人员需要穿戴工作服且应在指定安全区域操作,仅 2 人符合安全规定,检测为 safe_area, smock, safe,其他人员均在不同程度上有不安全行为。在图 5c 操作间 2 个场景下,涉及昏暗干扰。图 5c 定义了工作安全区域和危险区域,有 4 人符合安全规定,既穿工作服,又在安全区域活动,检测结果为 safe_area, smock, safe,还有 2 人检测为 dangerous_area, no_smock, dangerousII,该场景标签检测置信度均为 97%。

5 结 论

1) 通过优化 YOLOv5 算法结构,实现对复杂工业现场 12 类不安全行为的检测。文中模型具有实时性、多类行为识别和复杂环境适应性等特点,为工业现场的安全管理提供强有力的技术支持和保障。

2) 复杂环境下的不安全行为检测需要结合设备状态和外部环境信息,未来可融合多源数据分析作业人员不安全行为。

3) 不安全行为涉及面广,文中检测模型针对工作服、安全帽和危险、安全区域进行检测,没有对人员摔倒、蹦跳等行为训练,以后需扩充模型数据量并结合姿态信息,识别更多的不安全行为。

参 考 文 献

- [1] FANTONI G, SANTOCHI M, DINI G, et al. Grasping devices and methods in automated production processes[J]. CIRP Annals, 2014, 63(2): 679-701.
- [2] GONZALEZ-DE-SANTOS P, FERNÁNDEZ R, SEPÚLVEDA D, et al. Field robots for intelligent farms: inhering features from industry[J]. Agronomy, 2020, 10(11): DOI:10.3390/agronomy10111638.
- [3] RAUSAND M, UTNE I. Product safety: principles and practices in a life cycle perspective[J]. Safety Science, 2009, 47(7): 939-947.
- [4] 刘林,梅强,常志朋. 国内 70 年来员工不安全行为研究: 发展阶段、研究热点及趋势分析[J]. 中国安全科学学报, 2021, 31(3): 1-12.
LIU Lin, MEI Qiang, CHANG Zhipeng. Research on employees' unsafe behavior in China from 1949 to 2018: development stage, research hotspot and trend analysis[J]. China Safety Science Journal, 2021, 31(3): 1-12.
- [5] 侯志强, 韩崇昭. 视觉跟踪技术综述[J]. 自动化学报, 2006, 32(4): 603-617.
HOU Zhiqiang, HAN Chongzhao. A survey of visual tracking[J]. Acta Automatica Sinica, 2006, 32(4): 603-617.
- [6] WEINSTEIN B G. A computer vision for animal ecology[J]. Journal of Animal Ecology, 2018, 87(3): 533-545.
- [7] XU Shuyuan, WANG Jun, SHOU Wenchi, et al. Computer vision techniques in construction: a critical review[J]. Archives of Computational Methods in Engineering, 2021, 28(1): 3383-3397.
- [8] JIANG Meng, CUI Peng, FALOUTSOS C. Suspicious behavior detection: current trends and future directions[J]. IEEE Intelligent Systems, 2016, 31(1): 31-39.
- [9] 范冰倩, 董秉聿, 王彪, 等. 基于深度学习的地铁施工作业人员不安全行为识别与应用[J], 中国安全科学学报 2023, 33(1): 41-47.
FAN Bingqian, DONG Bingyu, WANG Biao, et al. Identification and application of unsafe behaviors of subway construction workers based on deep learning[J]. China Safety Science Journal, 2023, 33(1): 41-47.
- [10] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005: 886-893.
- [11] HUNG P D, SU N T. Unsafe construction behavior classification using deep convolutional neural network[J]. Pattern Recognition and Image Analysis, 2021, 31(1): 271-284.
- [12] PATWAL A, DIWAKAR M, TRIPATH V, et al. An investigation of videos for abnormal behavior detection[J]. Procedia Computer Science, 2023, 218: 2264-2272.
- [13] JOSEPH R, SANTOSH D, ROSS G, et al. You only look once: unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [14] 朱张莉, 饶元, 吴渊, 等. 注意力机制在深度学习中的研究进展[J]. 中文信息学报, 2019, 33(6): 1-11.
ZHU Zhangli, RAO Yuan, WU Yuan, et al. Research progress of attention mechanism in deep learning[J]. Journal of Chinese Information Processing, 2019, 33(6): 1-11.
- [15] KIM J H, KIM N, PARK Y W, et al. Object detection and classification based on YOLO-V5 with improved maritime dataset[J]. Journal of Marine Science and Engineering, 2022, 10(3): 377-379.
- [16] LIU Zhuang, LI Jianguo, ZHANG Changshui, et al. Learning efficient convolutional networks through network slimming[C]. IEEE International Conference on Computer Vision, 2017: 2736-2744.

作者简介: 纪执安 (1994—), 男, 河北沧州人, 硕士, 工程师, 主要从事机器学习、智能检测和图像处理等方面的工作。E-mail: 748830478@qq.com。

