

基于深度学习的铸件缺陷检测

鲍春生^{1,2} 谢刚¹ 王银^{1,2} 谢新林¹

(1.太原科技大学电子信息工程学院; 2.太原科技大学先进控制与装备智能化山西省重点实验室)

摘要 提出一种 YOLOv3-Mv2 的检测模型,对铸件划痕缺陷进行检测。首先,选择更小的骨架网络 MobileNetv2 替换原来的特征提取网络结构 Darknet53,以减少网络的计算参数量,提升检测速度;其次,在深层特征和浅层特征相融合的基础上增加新的检测尺度,以增强对小缺陷目标的检测能力;然后将 BN(Batch Normalization)层的参数重新计算,与卷积层共用连通区域数据,加快模型前向推断的速度;引入 CIoU(Complete IoU)函数提高定位准确度。结果表明,YOLOv3-Mv2 与原 YOLOv3 算法相比,平均精度(mAP)值提高了 5.42%,实时性提高了 23 f/s。

关键词 缺陷检测; MobileNetv2; 合并参数; 检测尺度; 损失函数

中图分类号 TP311; TG249; TG245

文献标志码 A

DOI:10.15980/j.tzzz.2021.05.012

开放科学(资源服务)标识码(OSID):



Casting Defect Detection Based on Deep Learning

Bao Chunsheng^{1,2}, Xie Gang¹, Wang Yin^{1,2}, Xie Xinlin¹

(1.School of Electronic Information Engineering, Taiyuan University of Science and Technology; 2.Shanxi Key Laboratory of Advanced Control and Intelligent Equipment, Taiyuan University of Science and Technology)

Abstract: A test model of YOLOv3-Mv2 was proposed to detect the scratches and defects of castings. Firstly, a smaller skeleton network MobileNetv2 instead of the original feature extraction network structure Darknet53 was chosen to reduce the amounts of computational parameters of the network and to accelerate detection. Secondly, a new detection scale based on the fusion of deep and shallow features was added to enhance the detection capability of the small defect target. Then the parameters of the BN layer were recalculated, sharing connected area data with the reel layer to speed up the forward inference of the model. Finally, the Complete IoU function was introduced to improve positioning accuracy. The results indicate that the average accuracy (mAP) value and the real-time performance of YOLOv3-Mv2 were improved by 5.42% and 23 f/s than those of the original YOLOv3 algorithm, respectively.

Key Words: Defect Detection, MobileNetv2, Merge Parameters, Detection Scale, Loss Function

随着现代工业的发展,铸件品质保证尤为重要。对于表面有缺陷的铸件,采用人工检测方法存在效率低、不宜长时间操作,易造成误检等问题^[1]。传统机器视觉检测法^[2]和深度学习检测^[3]是目前常用的缺陷检测方法。

传统机器视觉检测法先采集图片,接着使用图像处理方法获得结果。胡浩等^[4]对 3 类不同的磁瓦缺陷特征设计了不同的检测方法。戴斌宇等^[5]提出一种基于连通域特征组合的干扰抑制算法,并利用图像差分法提取缺陷进行分析判断物体表面缺陷存在情况。这类算法的速度较快,针对特定缺陷的检测效果较好,但是受

现场环境和个人操作等因素影响,对缺陷特征提取的能力差。国内人工智能技术发展很快,缺陷检测也因此引入了深度学习的方法,检测效果明显提高,应用范围更加广泛。深度学习使用卷积神经网络提取特征并自主学习,利用训练后的权重检测数据判定是否有缺陷。深度学习对于缺陷检测的精度更高,检测速度更快^[6]。其目标检测方法主要是两类:一类是以 Faster-RCNN 为代表的 Two-stage 的检测方法^[7],这类算法的检测精度较高,但这类算法先生成候选框,然后再对每个候选框进行分类,所以检测速度较慢,影响实时性。还有一类是以 SSD、YOLO 为代表的 One-stage 检测方法^[8]。利

收稿日期:2020-10-08;修改稿收到日期:2020-11-09
基金项目:山西省重大科技专项资助项目(20191102009);山西省重点研发计划(国际合作)资助项目(201803D421039);山西省重点研发计划资助项目(201903D121130,201803D121025);山西省基础研究计划资助项目(201901D111265);山西省研究生教育创新项目(2020SY422)
第一作者简介:鲍春生,男,1995 年出生,硕士研究生,太原(030024),E-mail: bcs520525@163.com
通信作者:谢刚,男,1972 年出生,博士研究生,太原(030024),E-mail: xiegang@tyust.edu.cn
引用格式:鲍春生,谢刚,王银,等. 基于深度学习的铸件缺陷检测[J]. 特种铸造及有色合金, 2021, 41(5): 580-584.
BAO C S, XIE G, WANG Y, et al. Casting defect detection based on deep learning[J]. Special Casting & Nonferrous Alloys, 2021, 41(5): 580-584.

用回归思想一步计算出要检测的目标坐标并判断目标类别,这类方法检测速度大大提升,但精度略低于 Two-stage 方法。胡贵桂^[9]采用残差网络(Resnet101)结构取代原先的特征提取网络,增强了特征提取的效果。ZHANG J Q 等^[10]用两个密集的网络模块替代了两个残余网络模块结构,提升了精度。SANDLER M 等^[11]提出了一种与空间金字塔分解相结合的可视化模型,并借助密集连接结构,提高了精度。

本课题采用 YOLOv3-Mv2 网络,对不同大小的划痕缺陷位置进行检测,首先选择更小的网络 MobileNetv2 替换原来的特征提取网络结构 Darknet53,以减少网络的计算参数量,提升检测速度,接着添加一个网络预测尺度,以增强对小缺陷目标的检测能力,并将 BN 层参数合并到卷积层,加快推演的速度,并引入 CI-oU 函数提高定位准确度。

1 YOLOv3 的介绍

YOLOv3^[12]是在 YOLOv1、YOLOv2^[13]基础上经过改善得到的网络结构。和 Faster R-CNN 的方法不同,YOLOv3 使用回归方法,将一张图片分割成 $N \times N$ 个网格,可以同时检测多目标,速度也较快。YOLOv3 的特征提取主干网络是 Darknet53,借鉴了残差网络 ResNet^[14]的思想,在层与层之间设置 shortcut(快捷链路),选择性地跳过部分卷积,解决了网络训练深层次时优化问题,加强了对特征的提取。其引入了特征金字塔思想,使用了 3 个不同尺度检测目标。另外,不使用 Softmax 对每个框进行分类,因为 Softmax 使得每个框分配一个类别,对于检测的图像,目标可能有重叠的标签,而且 Softmax 可被独立的多个逻辑分类器替代,且准确率不下降。

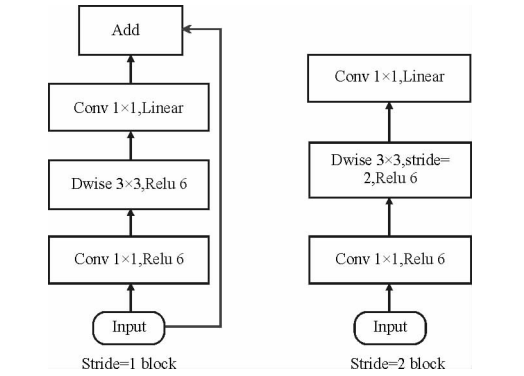
2 改进的 YOLOv3 算法

为了提升网络实时性,在保证精度的基础上,主要对网络做了如下改进。

2.1 改进骨干网络

YOLOv3 的结构是借鉴残差网络的思想设计的,在加强特征提取的同时,也增加了许多参数,一定程度上影响了实时性。所以利用 MobileNetv2 可以有效减少参数的同时保证一定的准确度。MobileNetv2 保留了 MobileNetv1 的深度可分离卷积等优点,并加入了具有线性瓶颈的逆残差结构(Inverted Residuals),与传统 ResNet 结构中先对 bottlenecks 降维再复原不同,这里先用一个 expand layer 进行升维,然后再进行通道卷积(DepthwiseConvolution),最后再 1×1 降维复原。提

出线性瓶颈(Linear Bottlenecks),在维度较小的 1×1 卷积之后不再采用 ReLU 操作,避免对特征的破坏。MobileNetv2 的基础结构图见图 1。更换骨干网络后检测对比见表 1。



(a) 步长为 1 逆残差结构图 (b) 步长为 2 逆残差结构图

图 1 MobileNetv2 基础图

Fig.1 Basic chart of MobileNet2

表 1 更换骨干网络后对比

Tab.1 Comparison before and after replacing the backbone network

	检测速度/(f · s ⁻¹)	检测精度/%
改进前	22	90.14
改进后	40	90.02

2.2 增加预测尺度

为了加强对小目标的检测,YOLOv3 网络采用了 3 个不同尺度的特征图来进行对象检测。网络的最终输出有 3 个尺度,分别为 $1/32$ 、 $1/16$ 、 $1/8$;在第 79 层之后经过几个卷积操作得到的是 $1/32$ 的预测结果,适合检测大目标。然后这个结果通过上采样与第 61 层的结果进行拼接,再经过几个卷积操作得到 $1/16$ 的预测结果;其具有中等尺度的感受野,适合检测中等尺度的对象。91 层的结果经过上采样之后再与第 36 层的结果进行拼接,经过几个卷积操作之后得到的是 $1/8$ 的结果,它的感受野最小,适合检测小尺寸的对象。通过上述操作方式,生成了 3 种预测尺度,即 13×13 、 26×26 、 52×52 ,可以对不同大小目标进行检测。针对铸件缺陷存在更小的缺陷,在原来 3 种检测的尺度上增加第 4 种检测尺度,在 52×52 的特征图上进行上采样,与 104×104 的特征图连接。改进后的结构见图 2。

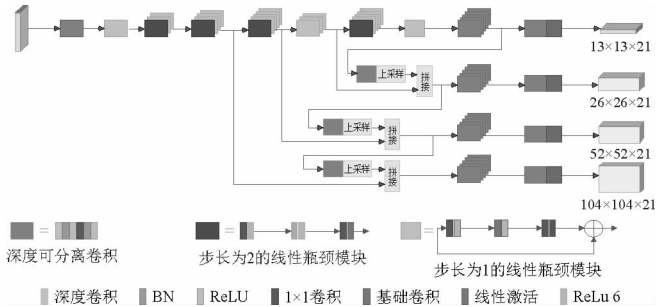


图 2 修改后的网络结构

Fig.2 The modified network structure

2.3 合并 BN 层与卷积层参数

BN 层可以加速网络的收敛,并且可以控制网络过拟合,但与此同时增加了网络前向推理时间,影响了性能,占用了更多的显存。所以将 BN 层参数合并到卷积层,提高模型前向推理速度。

卷积层运算:

$$Y_{\text{conv}} = W_{\text{conv}} \times X + b_{\text{conv}} \tag{1}$$

式中, W_{conv} 为卷积权重; b_{conv} 为卷积偏置; Y_{conv} 为卷积层的输出。

BN 层中:

$$Y_{\text{BN}} = a_{\text{bn}} \cdot \frac{x - m}{\sqrt{\delta + \epsilon}} + b_{\text{bn}} \tag{2}$$

将式(1)的输出作为 BN 层的输入代入式(2):

$$Y_{\text{BN}} = a_{\text{bn}} \cdot \frac{W_{\text{conv}} \times X + b_{\text{conv}} - m}{\sqrt{\delta + \epsilon}} + b_{\text{bn}} = \frac{a_{\text{bn}} \times W_{\text{conv}}}{\sqrt{\delta + \epsilon}} \times X + \frac{a_{\text{bn}}(b_{\text{conv}} - m)}{\sqrt{\delta + \epsilon}} + b_{\text{bn}} \tag{3}$$

$$W'_{\text{conv}} = \frac{a_{\text{bn}} \times W_{\text{conv}}}{\sqrt{\delta + \epsilon}} \tag{4}$$

$$b'_{\text{conv}} = \frac{a_{\text{bn}}(b_{\text{conv}} - m)}{\sqrt{\delta + \epsilon}} + b_{\text{bn}} \tag{5}$$

式中, X 是卷积层的输入; Y_{BN} 是 BN 层的输出; a_{bn} 和 b_{bn} 为 BN 层的网络参数; m 为均值; δ 为方差; ϵ 为一个较小量,防止分母为 0; W'_{conv} 为合并后卷积层权重; b'_{conv} 为合并后卷积层偏置。合并后共同使用 Blob 数据,减少显存使用,加快了推理速度,检测速度与精度对比见表 2。

表 2 合并参数后对比

Tab.2 Parameter comparison before and after combination		
	检测速度/(f · s ⁻¹)	检测精度/%
改进前	22	90.14
改进后	26	90.21

2.4 损失函数的改进

IoU(Intersection over union)是当前目标检测算法性能常见评判标准之一,它是预测框与真实框的交并比,可表示为:

$$\text{IoU} = \frac{A \cap B}{A \cup B} \tag{6}$$

式中, A 为预测框; B 为真实框。IoU 虽然可以一定程度上反映预测框和真实框的检测效果,但是当预测框和真实框没有交集时就不适用了,于是引入 CIoU(Complete intersection over union),CIoU 解决了一般 IoU 无法直接优化两个框没有重叠的部分,CIoU 考虑了目标与锚点框(anchor)之间的距离、尺度、惩罚项以及重叠率,使目标框回归变得更加稳定,避免 IoU 和 GIoU

(Generalized intersection over union)一样在训练过程中出现发散等问题,而惩罚因子考虑了预测框长宽比拟合目标框的长宽比。CIoU 的公式为:

$$\text{CIoU} = \text{IoU} - \frac{\rho^2(b, b^{\text{gt}})}{c^2} - \alpha\beta \tag{7}$$

式中, $\rho^2(b, b^{\text{gt}})$ 为预测框和真实框的中心点的欧式距离; c 为同时包含预测框和真实框的最小闭包区域的对角线距离; α 为权衡参数; β 衡量了长宽比的一致性,其中 α, β 可表述如下:

$$\alpha = \frac{\beta}{1 - \text{IoU} + \beta} \tag{8}$$

$$\beta = \frac{4}{\pi^2} \left(\arctan \frac{w_{\text{gt}}}{h_{\text{gt}}} - \arctan \frac{w}{h} \right)^2 \tag{9}$$

对应的损失函数为:

$$\text{LOSS}_{\text{IoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha\beta \tag{10}$$

2.5 模型的检测过程

Input:

Step 1: 输入图片,对图片代销进行调节;

Step 2: 经过 3×3 的逐通道卷积和 1×1 的逐点卷积,进行特征提取并减少参数数量的计算;

Step 3: 经过步长为 1 的线性瓶颈模块;

Step 4: 经过步长为 2 的线性瓶颈模块和步长为 1 的线性瓶颈模块,重复此步骤;

Step 5: 将多个尺度的特征进行融合,输出目标检测结果。

Output

3 试验结果与分析

3.1 试验平台

在 Ubuntu16.04 操作系统下进行试验,计算机 CPU 为 i7-9750H,内存为 16 G, GPU 为 GTX1660ti,显存为 6 GB,python 版本为 3.6.5,安装了 Cuda10.0 和 Cudnn7.5.1,加速 GPU 的计算与运行,同时安装了 Opencv4.1 和 Tensorflow1.13.1 以支持网络的训练和检测。

3.2 数据集

数据集共 2 000 张图片,由合作工厂提供,其中 1 530 张有缺陷,470 张无缺陷。训练集和测试集的比例按照 9 : 1,前 1 800 张作为训练集,后 200 张作为测试集。利用 labelImg 软件对图片进行人工标注,将图片缺陷的位置使用矩形框框出来,并标上缺陷的标签,该软件标注后会生成.xml 文件,再用 python 编写代码转换成.txt 文件。试验部分数据集如下:本试验数据集

为铸件缺陷数据集,主要为划痕缺陷,缺陷是狭长的,长短不一,有些划痕比较大,有些划痕比较小。

3.3 模型的训练

对改进的 YOLOv3 网络进行训练时训练集与验证集的比为 9 : 1,对网络参数进行初始化,学习率设为 0.001,迭代次数设为 12 000,批量大小根据电脑性能调为 8,衰减系数设为 0.000 5。训练的 loss 和 AvgIOU 图分别见图 3 和图 4。

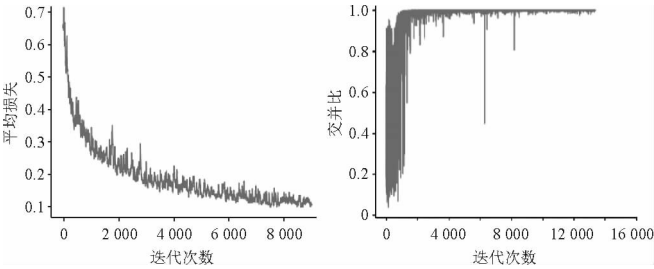


图 3 平均损失函数曲线
Fig.3 Curves of average loss function

图 4 平均交并比
Fig.4 Average cross-over ratio

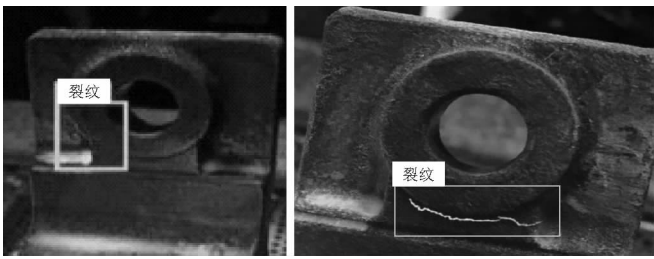
3.4 模型效果测试

利用训练 12 000 次得到的权重文件对测试集进行测试,见图 5。



图 5 小缺陷检测效果
Fig.5 Detection effect of small defects

数据集主要是铸件的划痕缺陷,划痕的大小及位置各不相同,数据集中的划痕只有一个,大划痕和小划痕的缺陷检测效果都很好,见图 6,大划痕的缺陷检测精度可以达到 98.34%,小划痕的缺陷检测精度达到 94.78%,可以有效检测出划痕缺陷的位置信息。



(a) 小缺陷检测图 (b) 大缺陷检测图
图 6 大、小缺陷检测图

Fig.6 Detection figures of small and large defects

3.5 测试对比结果与分析

为验证检测方法的有效性,将改进后的算法与原网络进行性能对比(相同试验条件下),具体评价指标为平均检测精度(mAP)、召回率(recall)、平均 IOU 以及检测速率。

表 3 测试对比结果

Tab.3 The comparative results of the test				
模型	平均检测精度/%	召回率/%	平均 IOU/%	检测速率/(f · s ⁻¹)
原网络	91.14	86.25	73.34	22
改进后	96.56	91.34	76.26	45

从表 3 可知,相对原网络,mAP 值提升 5.95%,检测速度提升了 23 f/s,召回率和平均 IOU 也有所提高,说明本方法具有较好效果。在相同试验条件下,与其他网络对比结果见表 4。可以看出,原 YOLOv3 算法的检测速度比 two-stage 算法要快很多,但是检测精度不如 two-stage 算法,本课题提出的 YOLOv3-Mv2 算法在检测速度与检测精度上都要比当前常用的目标检测算法高。

表 4 算法对比结果

Tab.4 The comparative results of algorithm					
模型	基础	尺寸/nm	IOU	mAP/%	检测速率/(f · s ⁻¹)
FasterR-CNN	VGG16	600×600	0.82	93.21	7
SSD	VGG16	300×300	—	87.23	26
YOLOv3	Darknet-53	416×416	0.73	91.14	35
YOLOv3-Mv2	MobileNetv2	416×416	0.76	96.56	45

4 结 语

提出一种 YOLOv3-Mv2 网络模型对工业生产中的铸件进行缺陷检测,首先选择更小的网络(MobileNetv2)替换原来的特征提取网络结构 Darknet53,减少网络的计算参数量,提升检测速度,接着添加一个网络预测尺度,以增强对小缺陷目标的检测能力;然后合并 BN 层的参数到卷积层,加快推演的速度;最后引用 CI-oU 增加定位准确度。从结果可知,提出的 YOLOv3-Mv2 网络模型在检测精度和检测速度上都有一定的提高。后续将继续深层次优化模型算法,将网络移植到移动设备上。

参 考 文 献

[1] 李少波,杨静,王铮,等.缺陷检测技术的发展与应用研究综述[J].自动化学报,2020,46(11): 2 319-2 336.
[2] JIAN C X, GAO J, AO Y H, et al. Automatic surface defect detection for mobile phone screen glass based on machine vision[J]. Applied Soft Computing, 2017, 52(52): 348-358.
[3] GAIDHANE V H, HOTE Y V, SINGH V. An efficient similarity measure approach for PCB surface defect detection[J]. Pattern Analysis and Applications, 2018, 21(1): 277-289.

基于 Creo 的铸造浇注系统参数化自动建模技术

刘颖卓¹ 崔恩强¹ 崔宇翔² 邹文兵¹

(1.上海航天精密机械研究所; 2.合肥工业大学材料科学与工程学院)

摘 要 针对建模复杂的零件浇注系统,在 Creo 环境下基于参数化自动建模设计进行二次开发。以族表技术为依托,并结合 Pro/TOOLKIT 进行二次开发,介绍了进行参数化自动建模的关键技术和开发步骤,弥补了通过族表自动建模的自由性和交互性差的缺陷。通过上述参数化自动建模,大大提高了建模速度,从整体上提高了设计效率。

关键词 参数化;族表;二次开发;Pro/TOOLKIT

中图分类号 TP311;TG249 **文献标志码** A

DOI:10.15980/j.tzzz.2021.05.013

开放科学(资源服务)标识码(OSID):



Automatic Modeling Technology of Casting Gating System Parameterization
Based on Creo Parametric

Liu Yingzhuo¹, Cui Enqiang¹, Cui Yuxiang², Zou Wenbing¹

(1.Shanghai Spaceflight Precision Machinery Institute;

2.School of Material Science and Engineering, Hefei University of Technology)

Abstract: In view of complex modeling gating system, the parametric automatic modeling design was performed in the Creo environment for secondary development. Based on family table technology and combined with Pro/TOOLKIT for secondary development, the key technologies and development steps of parametric automatic modeling were introduced, making up for the defects of freedom and poor interaction of automatic modeling through family table. The modeling speed was greatly improved, and the design efficiency was increased overall via the above parametric automatic modeling.

Key Words: Parameterization, Family Table, Secondary Development, Pro/TOOLKIT

铸造浇注系统对铸件品质起着关键作用。在浇注系统设计时,由于浇注系统结构形式复杂,三维建模往往需要花费较长时间和精力,影响工艺设计的效率。Creo 软件是一款以参数化技术为特征的通用化程

收稿日期:2020-10-15;**修改稿收到日期:**2020-11-28
第一作者简介:刘颖卓,男,1987 年出生,高级工程师,上海(201600),E-mail:15221009513@163.com
引用格式:刘颖卓,崔恩强,崔宇翔,等. 基于 Creo 的铸造浇注系统参数化自动建模技术[J]. 特种铸造及有色合金,2021,41(5):584-587.
LIU Y Z, CUI E Q, CUI Y X, et al. Automatic modeling technology of casting gating system parameterization based on creo parametric[J]. Special Casting & Nonferrous Alloys, 2021, 41(5):584-587.

[4] 胡浩,李俊峰,沈军民.基于机器视觉的小磁瓦表面微缺陷检测方法研究[J].机电工程,2019,36(2):117-123,184.

[5] 戴斌宇,吴静静.基于机器视觉的定子表面缺陷检测系统研究[J].传感技术学报,2019,32(10):1589-1594.

[6] GUO Y, OERLEMANS A, et al. Deep learning for visual understanding[J]. Neurocomputing, 2016, C187: 27-48.

[7] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN:Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[8] FU C Y, LIU W, RANG A, et al. DSSD:Deconvolutional single shot detector[J]. Computer Science, 2017, 1(1): 10-21.

[9] 胡贵桂.改进的 YOLOv3 算法在道路环境目标检测中的应用[J].汽车实用技术,2020(5):117-121.

[10] ZHANG J Q, KANG X, NI H J, et al. Surface defect detection of steel strips based on classification priority YOLOv3-dense network [J]. Ironmaking & Steelmaking, 2020, 13(7): 1-12.

[11] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks[J].IEEE Access, 2018, 5: 1-12.

[12] JING J F, ZHUO D, ZHANG H H, et al. Fabric defect detection using the improved YOLOv3 model[J]. Journal of Engineered Fibers and Fabrics, 2020, 15: 1-10.

[13] ZHANG J M, HUANG M T, JIN X K, et al. A real-time Chinese traffic sign detection algorithm based on modified YOLOv2 [J]. Algorithms, 2017, 10(4): 121-127.

[14] WU Z F, SHEN C H, HENGEL A V D. Wider or deeper: Revisiting the resnet model for visual recognition[J]. Pattern Recognition, 2019, 90: 119-133.

(编辑:刘晨辉)