

DOI:10.19651/j.cnki.emt.2108495

基于改进的YOLOv4彩色数字仪表读数识别方法*

侯卓成 欧阳华 胡鑫 尹洋
(海军工程大学电气工程学院 武汉 430033)

摘要: 针对传统数字仪表识别方法计算量大、实时性不够和精度较低的问题,研究了一种深度学习与图像处理相结合的识别方法。为减少计算量,在YOLOv4网络中引入GhostNet作为基础网络,同时在YOLOv4非主干网络中引入深度可分离卷积以及Ghost模块来减少参数,并使用h-swish激活函数提高精度。为了凸显二值化过程中彩色仪表的色彩信息,在数字提取过程,研究了一种基于彩色模型多阈值分割的数字二值化方法,对RGB图像的色彩主色进行增强,然后转化为HSI图像,并且通过多阈值处理将满足条件的像素点保留,从而得到二值化图像,相比于传统图像预处理算法可以更好地提取数字信息。实验结果表明改进的识别方法在测试集上准确率达到87.98% mAP,检测速度提高到37.2 FPS,在数字仪表定位识别中效果显著。

关键词: YOLOv4; GhostNet; 数字识别; 仪表检测

中图分类号: TP39 **文献标识码:** A **国家标准学科分类代码:** 510.1050

Improved YOLOv4 color digital instrument reading recognition method

Hou Zhuocheng Ouyang Hua Hu Xin Yin Yang
(College of Electrical Engineering, Naval University of Engineering, Wuhan 430033, China)

Abstract: The traditional digital instrument recognition method has a large amount of computational amount, not enough real-time, low accuracy. This paper studies a meter identification method is studied in combination with deep learning and image processing. In order to reduce the amount of computation, YOLOv4 target detection network is used and GhostNet is adopted as YOLOv4 basic network. At the same time, the depthwise separable convolution and Ghost module can be introduced in YOLOv4 to reduce the amount of parameters, and the h-swish activation function is applied to increase the accuracy. In order to highlight color information in the image binarization process, a digital binarized method is studied based on color model multi-threshold segmentation. The main color of RGB image is enhanced, and then converts to an HSI image, and then the pixel point satisfying the condition will be reserved by multi-threshold processing, thereby obtaining a binarized image. Digital information can be better extracted in comparison with traditional image pretreatment algorithms. Experimental results show that the proposed method of reaches 87.98% mAP on the test data set, and detection speed is increased to 37.2 FPS, and then the effect is significant in digital instrument positioning and digital inspection.

Keywords: YOLOv4; GhostNet; digital inspection; instrument inspection

0 引言

针对数字式仪表的识别,主要分为传统图像处理的目标检测方法和基于深度学习的目标检测方法。在传统图像处理方面,杨肖^[1]利用HSV(hue, saturation, value)空间的H分量的信息来定位出待识别的数字符号完成对字符的分割。张艳玲等^[2]利用亮度信息提取数字区域得到较好的效果。陈家翔等^[3]根据数字仪表字符的特点,将没有点亮的数码管作为特征,在图像上定义7个子区域来对应7个

数码管进行数字识别,使得识别时间不超过0.5 ms,并且有一定的抗干扰能力。周晓彦等^[4]概述了基于深度学习检测方面,近些年来,随着各种目标检测网络包括RCNN(regions with CNN features)系列算法^[5-7]、YOLO(you only look once)系列算法^[8]和SSD(single shot multibox detector)算法^[9]等相继提出,通过深度学习算法进行仪表数字区域的定位和识别,提高对数字式仪表的识别准确率也是目前目标检测识别算法研究的热门方向。蔡梦倩

收稿日期:2021-12-02

* 基金项目:极区海图基准和导航误差控制理论及其应用研究(41876222)项目资助

等^[10]利用全卷积神经网络来实现对数字式仪表读数的识别,通过图到图的预测来获取最后的结果。朱立倩^[11]利用改进的Faster R-CNN进行数显仪表的自动识别,融合注意力机制,提高了分类及定位的准确性。孙顺远等^[12]采用改进的SSD算法对仪表图像进行检测,提高了对仪表的检测能力。但深度学习这类算法往往结构复杂,运算量大,在运算能力较差的设备上难以部署使用,因此需要使用轻量级神经网络来完成数字检测任务。YOLOv4^[13]是YOLO系列目标检测网络算法中的第4个版本,检测精度和检测速度都有所提高。解尧婷等^[14]、彭继慎等^[15]、张明路等^[16]分别对YOLOv4算法进行轻量化改进,在不同程度上提高了YOLOv4算法的性能。

为加快仪表识别运算速度,提高识别准确率,本文通过仪表定位、图像二值化、数字分割以及数字识别4个部分完成对彩色数字仪表识别。为了减少计算量,本文研究了改进的YOLOv4网络来完成目标检测任务。为了更准确地提取彩色数字信息,提出了一种基于RGB图像以及HSI图像两种图像显示模式上进行多阈值处理的二值化方法。采用双向投影法^[17],将字符进行逐个分割。使用更加方便快捷的模板匹配算法^[18]得到识别结果。最后,在智能巡检机器人装置上验证本文算法的有效性。

1 基于改进的YOLOv4算法的表盘定位

传统YOLOv4目标检测网络的主干提取网络使用了CSPDarkNet53,相比于YOLOv3算法,网络结构没有较大改变,但增加了特征金字塔模块、Mosaic数据增强、Label Smoothing平滑、CIoU、学习率余弦退火衰减、Mish激活函数等方法,从而提高了检测精度。

为减少网络计算量,本文选择Han等^[19]提出的GhostNet作为YOLOv4的主干网络。网络中Ghost模块的使用可以在大大减少参数量的同时生成同样的特征。Ghost模块结构如图1所示,对于输入的特征层,先使用普通卷积操作生成部分真实的特征层,然后在对每个特征层的通道使用深度可分离卷积处理得到Ghost特征层,最后将第1次卷积得到的真实特征层与深度可分离卷积^[20]得到的特征层堆叠在一起得到完整的特征层。

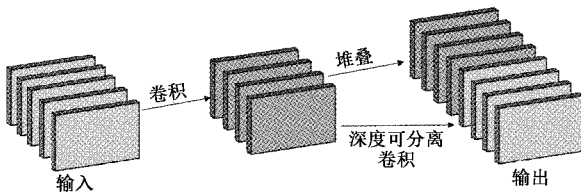


图1 Ghost模块结构

假设输入特征图大小为 $h_1 \times h_1 \times n$,输出特征图大小为 $h_2 \times h_2 \times m$,卷积核的尺寸为 $k \times k$,可知普通卷积的计算量为 C_1 :

$$C_1 = h_2 \times h_2 \times n \times k \times k \times m \quad (1)$$

而Ghost模块卷积计算分为两步:第1部分卷积核尺寸为 k_1 ,第2部分卷积核尺寸为 k_2 ,假设要生成 s 个幻影特征图过渡,则其计算量为 C_2 :

$$C_2 = h_2 \times h_2 \times \frac{m}{s} \times k_1 \times k_2 \times n \times (s-1) \times h_2 \times h_2 \times \frac{m}{s} \times k_2 \times k_2 \quad (2)$$

由式(1)、(2)可知,Ghost模块可以给模型带来压缩比为 s 的压缩,大大减少了网络计算量。

GhostNet主要由一系列Ghost bottleneck组成,其中Ghost bottleneck以Ghost模块为构建基础。GhostNet第1层是具有16个卷积核的标准卷积层,然后经过一系列Ghost bottleneck,通道逐渐增加。这些Ghost bottleneck根据其输入特征图的大小分为不同的阶段。本文使用到的GhostNet网络部分参数如表1所示。

表1 本文使用到的GhostNet网络部分参数表

输入	卷积	升维	输出通道	SE	步长
416×416×3	Conv2d	—	16	—	2
208×208×16	G-bneck	16	16	—	1
208×208×16	G-bneck	48	24	—	2
104×104×24	G-bneck	72	24	—	1
104×104×24	G-bneck	72	40	1	2
52×52×40	G-bneck	120	40	1	1
52×52×40	G-bneck	240	80	—	2
26×26×80	G-bneck	200	80	—	1
26×26×80	G-bneck	184	80	—	1
26×26×80	G-bneck	184	80	—	1
26×26×80	G-bneck	480	112	1	1
26×26×112	G-bneck	672	112	1	1
26×26×112	G-bneck	672	160	1	2
13×13×160	G-bneck	960	160	—	1
13×13×160	G-bneck	960	160	1	1
13×13×160	G-bneck	960	160	—	1
13×13×160	G-bneck	960	160	1	1

在网络中除了每个阶段的最后一个Ghost Bottleneck是stride=2,其他所有Ghost bottleneck都以stride=1进行应用。Ghost Bottlenecks结构如图2所示。

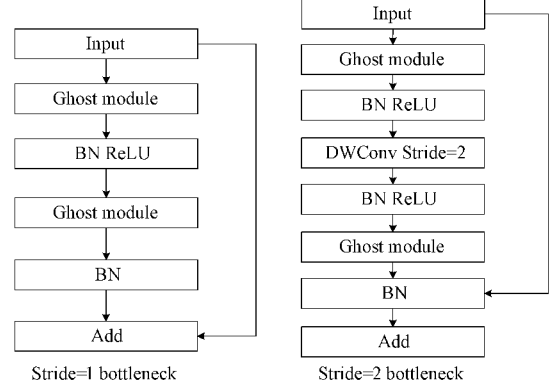


图2 不同步长的Ghost Bottlenecks结构

为了进一步减少计算量,本文在 YOLOv4 中的 PANet 结构中,使用 Ghost 模块进行卷积提取特征,在 YOLOv4 中的 YOLO Head 结构中将普通卷积 3×3 卷积进行替换为深度可分离卷积结构,以此来减少网络的参数量。深度

可分离卷积结构如图 3 所示,此卷积是用 3×3 大小的卷积核分别遍历所有通道的数据,得到了相应数量的特征图,接着用 1×1 大小的卷积核遍历每个特征图谱。深度可分离卷积的使用可以大大减少模型的参数量。

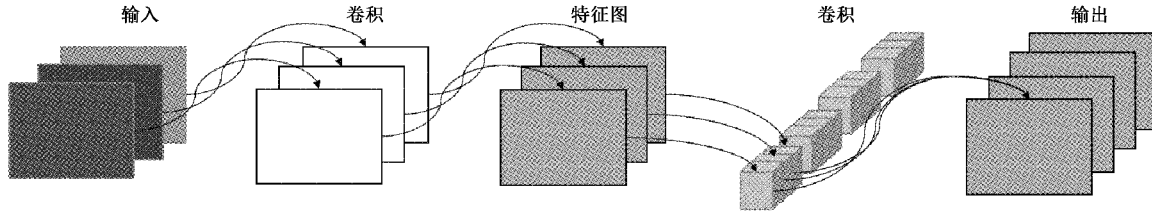


图 3 深度可分离卷积

为了确保识别精度的稳定性,在 YOLOv4 非主干网络部分以及深度可分离卷积结构块中引入 h-swish 激活函数^[21]。swish 激活函数在模型效果上优于 ReLU 激活函数,但是计算代价很大。而 h-swish 激活函数是用一个近似函数来逼近 swish 激活函数,相比于 swish 激活函数大大减少了计算量,本文使用 h-swish 激活函数可以在保证计算量的同时,提高其精度。3 种激活函数如式(3)~(5)所示,3 种激活函数曲线对比如图 4 所示。

$$\text{h-swish}[x] = x \frac{\text{ReLU6}(x + 3)}{6} \quad (3)$$

$$\text{swish}[x] = x \cdot \text{sigmoid}(\beta x) \quad (4)$$

$$\text{ReLU6}(x) = \min(\max(0, x), 6) \quad (5)$$

改进后的 GhostNet-YOLOv4 其整体结构如图 5 所示,分别在进行第 5 次、第 11 次、第 16 次 Ghost bottleneck 后进行提取,获得主干网络的 3 个初步有效特征层大小分

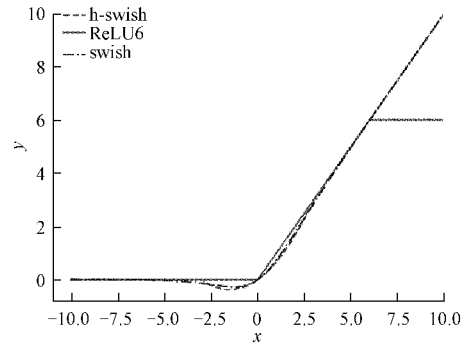


图 4 ReLU6、swish 和 h-swish 激活函数对比图

别为 $52 \times 52 \times 40$ 、 $26 \times 26 \times 112$ 、 $13 \times 13 \times 160$ 。而后利用加强特征提取网络 SPP 和 PANet 对初步获得的有效特征层进行特征融合,最后通过预测网络 YOLO Head 获得预测结果。

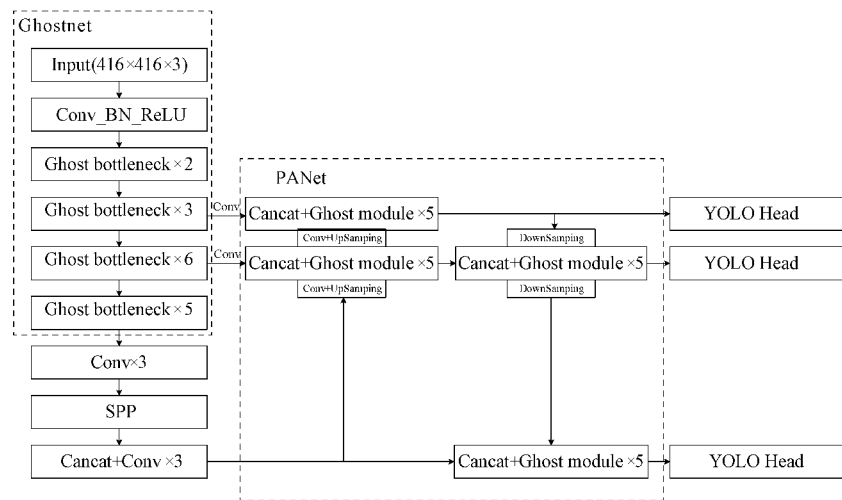


图 5 GhostNet-YOLOv4 结构

将 YOLOv4、GhostNet-YOLOv4 以及改进的 GhostNet-YOLOv4 的网络参数量进行对比,由图 6 可知,改进的 GhostNet-YOLOv4 参数量仅为 YOLOv4 的 15.7%,大大地减少了计算量,可以更快的检测数字仪表盘位置。

为了验证改进后的检测效果,本文使用操作系统为 Windows 10,Pytorch(1.8.0)框架,CPU 为 AMD Ryzen 7-4800H,GPU 为 Nvidia RTX2060,Cuda 版本为 11.2,对进行网络结构的修改实验。使用 voc 数据集格式对 820 张数字仪表盘图像进行标注训练,在 150 张数字仪表盘图像上进

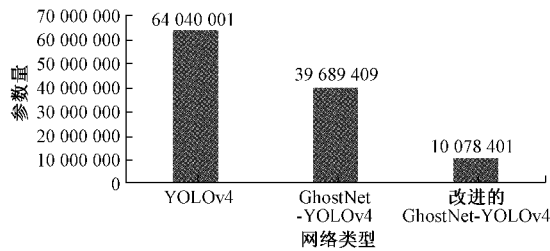


图6 改进前后网络参数量对比

行检测,训练过程为100个epoch。训练结束后对YOLOv4、GhostNet-YOLOv4和改进的GhostNet-YOLOv4的各项性能指标进行了对比。用平均精度均值(mean average precision, mAP)值来衡量检测精度,准确率(Precision)表示正确预测占全部预测为总的比例;召回率(Recall)表示正确预测为正占实际为总的比例;F1 Score又称为平衡F分数(balanced F Score)表示Precision和Recall的调和均值;FPS(frames per second)表示检测速度;Size表示权重的大小。

表2 改进前后各项性能指标的对比

性能指标	YOLOv4	GhostNet-YOLOv4	改进的 GhostNet-YOLOv4
平均精度均值/%	87.63	86.15	87.98
准确率/%	83.06	84.00	84.47
召回率/%	86.55	70.59	73.11
平衡F分数	0.85	0.77	0.78
FPS/(帧·s ⁻¹)	19.10	30.91	37.20
Size/M	244.4	150.2	37.3

由表2可知,改进后的GhostNet-YOLOv4可以在检测精度几乎维持不变的情况下,使模型权重大小下降到只有YOLOv4模型权重的15.3%,FPS提高到YOLOv4的1.95倍,检测速度大大提高,更适合在嵌入式设备上部署。

改进的GhostNet-YOLOv4网络检测效果如图7所示。



图7 GhostNet-YOLOv4网络测试结果

2 读数识别

2.1 数字提取

HSI(hue, saturation, intensity)颜色模型用H、S、I三个参数描述颜色特性,其中H表示颜色的波长,称为色调;S表示颜色的深浅程度,称为饱和度;I表示强度或亮度。HSI模型中I分量与图像的彩色信息无关,而H和S分量与人感受颜色的方式是紧密相联的。这些特点使得HSI模型非常适合彩色特性检测与分析。HSI模型反映了人感知颜色的基本属性,与人感知颜色的结果一一对应。因此,HSI模型被广泛应用。

彩色数字仪表中的数字往往是以红色或者绿色等主流颜色显示,在处理这类彩色仪表时,二值化过程大多采用全局阈值法、动态阈值法等传统方法,先将图像灰度处理,忽略了彩色图像的原本色彩信息,二值化效果往往不好,存在很多噪声影响。

为了更好地提取彩色数字,尽量只保留图像中的数字信息,本文设计了一种基于彩色模型多阈值分割的数字二值化方法,根据液晶显示屏数字仪表彩色数字显示原理,在RGB图像上根据彩色数字的主色Y(RGB分量中最大分量),设置阈值T,并根据阈值T进行判断,若主色Y大于阈值T将主色颜色增强至255。再将图像转换为HSI图像根据RGB转HSI图像计算公式,可知当主色Y被提高至255后I亮度分量将会增大,再根据亮度I,设置阈值IT。由于彩色仪表数字仪表的背景颜色大多为黑色或者白色,根据不同仪表颜色显示的差异判断阈值,阈值T选择在100~165之间,经过颜色增强处理后,阈值IT选取在0.3~0.5之间。为解决白色等背景色颜色过滤不充分对数字提取的影响,本文提出判别式如式(6)所示。

$$\frac{Y}{R+G+B} > IT \times \frac{R+G+B}{\sqrt{3} \times 255} \quad (6)$$

用来过滤掉白色等背景影响。图8所示为经判别式(6)处理前后二值化效果的差异,判别式(6)可以较好的过滤白色等背景信息。

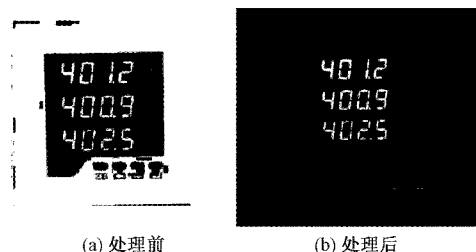


图8 经判别式(6)处理前后二值化效果对比图

通过多阈值处理将满足条件的像素点保留,得到二值化图像,再对二值化图像进行膨胀处理,避免数字离散化,使分割字符更加完整,最后再将图像进行腐蚀处理作为输出结果。

本文以实验室数字仪表图像为例进行分析。通过 RGB 转 HSI 图像公式(7)得到图像 I 分量。

$$H = \begin{cases} \theta, & G \geq B \\ 2\pi - \theta, & G < B \end{cases}$$

$$\theta = \cos^{-1} \left(\frac{(R - G) + (R - B)}{2 \sqrt{(R - G)^2 + (R - B)(G - B)}} \right) \quad (7)$$

$$S = 1 - \frac{3 \min(R, G, B)}{R + G + B}$$

$$I = \frac{R + G + B}{3}$$

选择阈值 T 为 135, 阈值 IT 为 0.3, 再根据判别式(8)进行二值化处理。

$$\begin{cases} R > G \& R > B \\ R > T \& I > IT \\ \frac{R}{R + G + B} > IT \times \frac{R + G + B}{\sqrt{3} \times 255} \end{cases} \quad (8)$$

为了比较本文算法所得二值化效果, 通过双阈值分割法、单阈值分割法、迭代法、大津法与本文提出的二值化方法进行对比如图 9 所示。

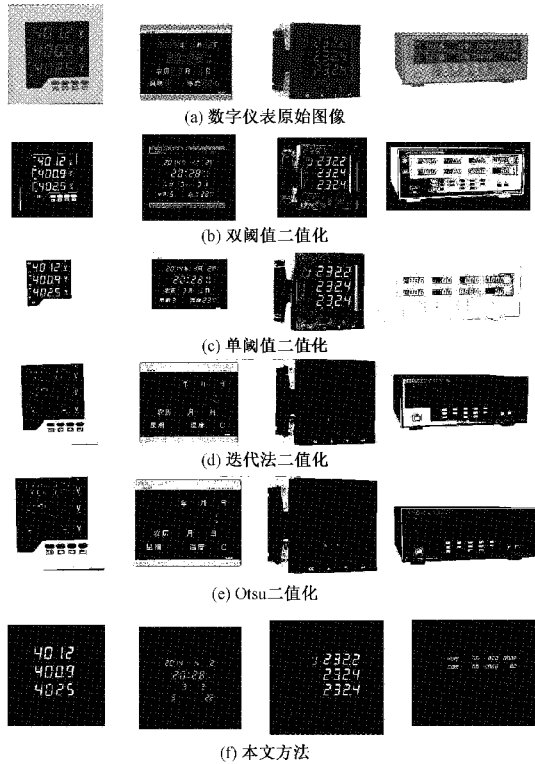


图 9 二值化方法对比图

由图 9 可知, 本文通过将彩色图像从 RGB 和 HSI 空间利用其颜色特性对其进行处理, 相对于传统二值化方法可以较完整地只保留彩色数字信息, 达到满足识别要求的数字提取效果。

2.2 数字分割

数字分割的好坏极大影响了读数识别的精度。经本文二

值化处理, 图片中只保留数字字符信息。为满足识别的需要, 本文将字符按双向投影法进行逐个分割, 对图像进行水平竖直投影得到数字的位置信息, 同时记录每一行或者每一列对应值的像素的个数, 然后根据这些个数确定数字的上下左右边界。其中像素的个数就像是一个阈值, 最后可以把每一行点的个数画出来便于直观地观察。利用二值化图像分布直方图进行分析, 从而找出相邻字符的分界点进行分割。由于单独的字符分割图像尺寸大小会有不同, 为了方便之后的字符识别, 将每一个分割字符进行尺寸的归一化处理, 本文为了数字识别更加方便准确本文将数字字符归一化为 36×36 大小的图片。数字分割结果如图 10 所示。



图 10 数字分割效果图

2.3 数字识别

LED 显示的数字仪表数字往往形式固定, 类型相对统一, 模板比较容易得到, 所以使用模板匹配算法往往更加方便快捷。首先选取合适的 0~9 这 10 个数字的模板, 将数字图像与模板图像进行处理归一化成大小相同的尺寸, 然后采用重合度和差别函数来度量字符和标准模板之间的相似程度并进行的对比, 相似度最大的则被认为是相同的数字, 从而确定了数字图像表示的值。该方法具有简单、易于实现等优点。数字识别结果如图 11 所示。



图 11 数字识别结果

3 实验验证

为了验证算法的有效性, 本文在智能巡检机器人实验装置如图 12 所示进行实验, 图中巡检机器人摄像头在吊轨上移动, 对数字仪表进行抓拍后, 传入给远程识别系统, 进行识别分析。



图 12 实施巡检图

数字式仪表的读数验证在于对数字显示区域的检测

和对屏幕内阿拉伯数字识别,巡检过程中针对数字仪表进行读数识别验证实验。

吊轨式智能巡检机器人对配电柜上的数字仪表进行巡检,通过在巡检平台进行实验,对数字式仪表进行检测识别,巡检30次,共检测300张仪表图像,数字3600个,仪表数字识别准确率达到98.94%,检测效果如图13所示。从彩色数字式仪表的识别过程中可以看出,利用本文所述算法应用到变电站巡检机器人中去进行巡检任务时,能够将捕获到的仪表示数进行准确识别,验证了本文算法的可行性。

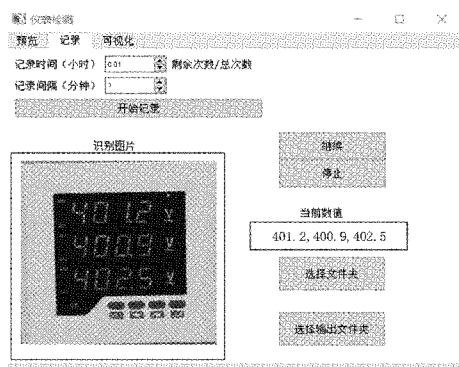


图13 数字式仪表识别界面

4 结 论

本文研究了一种改进的数字仪表识别算法,改进的GhostNet-YOLOv4网络模型可以在保证精度的前提下,使参数量仅有YOLOv4的15.7%,检测速度提高到37.2FPS,更加方便在嵌入式设备上使用。基于RGB以及HSI空间上的多阈值二值化方法,对于一些已知色彩分布的彩色数字仪表图像,能够过滤背景中其他差距比较大的颜色,可以较完整的只保留数字信息,达到符合识别要求的二值化效果。实验结果表明,改进算法的mAP在本文仪表测试集上达到87.98%,能够满足工业识别要求。下一步将对复杂背景下的仪表识别进行研究,设计出一种应用背景更加丰富的识别算法。

参考文献

- [1] 杨肖. 基于遗传算法的数字识别系统[D]. 广州:华南理工大学,2012.
- [2] 张艳玲,江仁焯,宁志刚,等. 彩色数字仪表表盘读数的自动识别系统[J]. 计算机测量与控制,2007(4):555-557.
- [3] 陈家翔,龙建忠,陶青川,等. 数字仪表显示值的快速识别方法[J]. 中国测试技术,2006(6):49-51,86.
- [4] 周晓彦,王珂,李凌燕. 基于深度学习的目标检测算法综述[J]. 电子测量技术,2017,40(11):89-93.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014:580-587.
- [6] GIRSHICK R. Fast R-CNN[C]. Proceedings of the

IEEE International Conference on Computer Vision, Los Alamitos: IEEE Computer Society Press, 2015: 1440-1448.

- [7] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards realtime object detection with region proposal networks [C]. Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015, 1: 91-99.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE International Conference on Computer Vision, Los Alamitos: IEEE Computer Society Press, 2016: 779-788.
- [9] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. 2015 14th European Conference on Computer Vision (ECCV), 2015: 1-17.
- [10] 蔡梦倩,张蕾,王炎,等. 基于全卷积网络的数字仪表字符识别方法[J]. 现代计算机(专业版),2018(2):38-43.
- [11] 朱立倩. 基于深度学习的数显仪表字符识别[J]. 计算机技术与发展,2020,30(6):141-144.
- [12] 孙顺远,杨挺. 基于深度学习的仪表口标检测算法[J]. 仪表技术与传感器,2021(6):104-108.
- [13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J/OL]. ArXiv, 2020, ArXiv: 2004. 10934.
- [14] 解尧婷,张丕状. 基于改进的YOLOv4输电线路小目标检测[J]. 国外电子测量技术,2021,40(2):47-51.
- [15] 彭继慎,孙礼鑫,王凯,等. 基于模型压缩的ED-YOLO电力巡检无人机避障口标检测算法[J]. 仪器仪表学报,2021,42(10):161-170.
- [16] 张明路,郭策,吕晓玲,等. 改进的轻量化YOLOv4用于电子元器件检测[J]. 电子测量与仪器学报,2021,35(10):17-23.
- [17] 周景超,陈锋,陈为多,等. 车牌字符分割的研究和实现[J]. 计算机工程,2006(5):238-240,243.
- [18] 魏武,张起森,王明俊,等. 一种基于模板匹配的车牌识别方法[J]. 中国公路学报,2001(1):106-108.
- [19] HAN K, WANG Y, TIAN Q, et al. GhostNet: More features from cheap operations[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2020: 1580-1589.
- [20] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv, 2017, ArXiv, 2017: 1704. 04861.
- [21] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. ICCV 2019: Proceedings of the IEEE International Conference on Computer Vision, Piscataway: IEEE, 2019: 1314-1324.

作者简介

- 侯卓成,硕士研究生,主要研究方向为计算机视觉。
E-mail: 2193513923@qq.com
- 欧阳华,副教授,主要研究方向为信号处理。
E-mail: 78ouyang@163.com
- 胡鑫,硕士研究生,主要研究方向为计算机视觉。
E-mail: 1810347286@qq.com
- 尹洋(通信作者),讲师,主要研究方向为控制工程。
E-mail: reeyan@163.com