

doi: 10.7690/bgzd.2025.05.012

基于深度神经网络的数字水表识别与应用

包霞, 褚燕华, 何月, 王丽颖, 王月明, 张晓琳

(内蒙古科技大学信息工程学院, 内蒙古 包头 014000)

摘要: 为提高水务公司对居民用水量的管理, 提出一种基于深度神经网络模型的数字式水表读数检测 and 识别方法。在原始 DBNet 网络模型的基础上对特征提取网络进行改进, 将特征提取网络 ResNet 网络替换成 DPN 双路径特征提取网络; 使用透视变换算法对倾斜的读数区域图像进行校正; 采用 YOLOv5s 网络模型对读数区域中的若干个数字进行识别; 通过水表识别应用系统将读数结果进行显示, 完成水表图像实时检测与识别的功能。实验结果表明: YOLOv5s 网络模型具有较好的鲁棒性, 能够准确识别出遮挡、模糊以及双半字符数字, 识别准确率高达 99.1%。

关键词: 读数识别; 深度神经网络; DBNet; 透视变换; 目标检测

中图分类号: TP391 **文献标志码:** A

Recognition and Application of Digital Water Meter Based on Deep Neural Network

Bao Xia, Chu Yanhua, He Yue, Wang Liying, Wang Yueming, Zhang Xiaolin

(School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014000, China)

Abstract: In order to improve the water company's management of residential water consumption, a digital water meter reading detection and recognition method based on deep neural network model is proposed. The feature extraction network is improved on the basis of the original DBNet network model, the feature extraction network ResNet network is replaced by the DPN double-path feature extraction network, the perspective transformation algorithm is used to correct the inclined reading area image, the YOLOv5s network model is used to recognize a plurality of digits in the reading area; The reading results are displayed by the water meter recognition application system, and the function of real-time detection and recognition of water meter images is completed. The experimental results show that the YOLOv5s network model has good robustness, and can accurately identify the occlusion, blur and double-half digits, and the recognition accuracy is as high as 99.1%.

Keywords: reading recognition; deep neural network; DBNet; perspective transformation; object detection

0 引言

数字式水表作为居民用水的度量工具, 在日常生活中随处可见。目前, 水表作为居民用水量的一个重要衡量标准, 水务公司采用客户端上传水表图像, 人工对水表图像进行读数, 易造成工作低效、费时、费力; 因此, 水表图像的自动读数是亟需解决的问题。

近年来, 随着图像识别技术与深度学习的不断发展, 水表读数识别已从传统的特征提取和模板匹配方法转变为基于卷积神经网络的端到端识别方法。汤鹏等^[1]使用灰度世界算法和霍夫变换算法等对图像进行预处理, 提高了数字识别准确度; 侯卓成等^[2]在 YOLOv4 网络模型中引入深度可分离卷积和 Ghost 模块, 提高了识别精度和减少了复杂度; 华泽玺等^[3]将 YOLOv4 与注意力机制结合构建一个通过有效通道注意该紧的深度可分离卷积, 提高了

变电站数字仪表检测的准确率和鲁棒性; 郭兰英等^[4]使用传统的图像处理技术与可变形卷积神经网络相结合, 提供了数字仪表识别的准确率以及网络的收敛速度; 周曼等^[5]通过光学字符识别 (optical character recognition, OCR) 提取数字仪表的数据, 然后利用 EAST 全卷积神经网络文字检测, 并使用 CNN-LSTM-CTC 文字识别, 提高了小数点的识别准确率; 综上所述, 将传统的图像处理方法与深度学习方法相结合识别水表读数已成为必然趋势。

为提高数字式水表读数识别的准确性和鲁棒性, 笔者基于深度学习方法提出自动水表读数的水表识别应用系统。本文中算法与文献[4]和[5]相比, 不再需要受噪音点以及字符个数的限制, 能直接检测读数区域中的数字, 提高了算法的鲁棒性识别效率。该算法通过定位水表图像中读数区域、矫正读数区域图像、识别读数区域中的数字, 采用水表识

收稿日期: 2024-08-10; 修回日期: 2024-09-15

基金项目: 内蒙古自治区自然科学基金(2020MS06008)

第一作者: 包霞(1999—), 女, 湖南人, 硕士。

别应用系统将水表图像的读数结果进行反馈。

1 读数区域检测

读数区域检测的整体流程如图 1 所示。整个流程包括读数区域定位和矫正 2 部分，其中读数区域定位通过改进的 DBNet 网络模型对原始输入图像进行训练，获取水表图像中读数区域的位置；读数区域矫正通过透视变换算法将读数区域定位获取的位置信息与目标位置信息进行映射，得到透视后的读数区域图像。

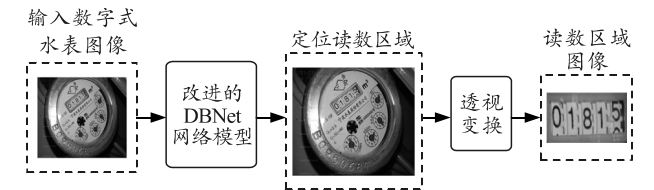


图 1 整体流程

1.1 读数区域定位

1.1.1 基于改进 DBNet 模型的读数区域定位

部分数据集如图 2 所示，该数据集来源于水务公司所管理的城镇居民的水表。由图 2 可以看出，由若干个数字组成的读数区域在水表图像中存在着多角度、遮挡等问题。笔者结合水表图像的特征，采用基于分割的 DBNet 文本行检测算法。由于 DBNet 特征提取模块采用的是 ResNet50 网络，ResNet 网络存在计算量大参数过多等问题，使得 DBNet 训练时间较长、性能降低；因此，笔者对 DBNet 的特征提取模块进行改进，将特征提取层的

ResNet 网络改为 DPN92^[6]网络。DPN92 网络将 ResNet 特征的再利用、低冗余性等与 DenseNet 的新特征生成等结合起来，有效解决了梯度消失或爆炸的问题，同时使得网络的复杂度降低、参数量减少。

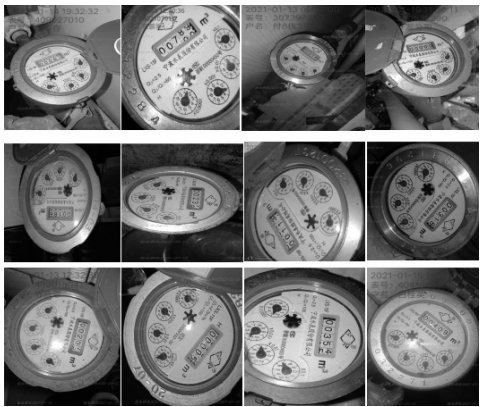


图 2 部分数据集

改进 DBNet 网络模型结构如图 3 所示，具体流程为：1) 输入图像经过 1 个 7*7 的卷积操作和步长为 2 的最大池化操作；2) 将输入图像投放到特征提取层的 DPN 特征提取网络，其中过渡层(Transition)由 BN+卷积层+池化层组成，用于每个 Block 之间的连接，可让下层的 Block 得到上层 Block 的输入和输出信息，使得数据的冗余和参数量减少^[7-8]；3) 经过 1*1 的卷积进行连接操作得到特征图 F，通过将特征图 F 送入 DBNet 网络进行预测得到概率图 P 和阈值图；4) 利用概率图 P 和特征图 T 进行可微分二值化 DB 计算，得到近似二值图，再经过 Box formulation 获取最终的图像检测结果。

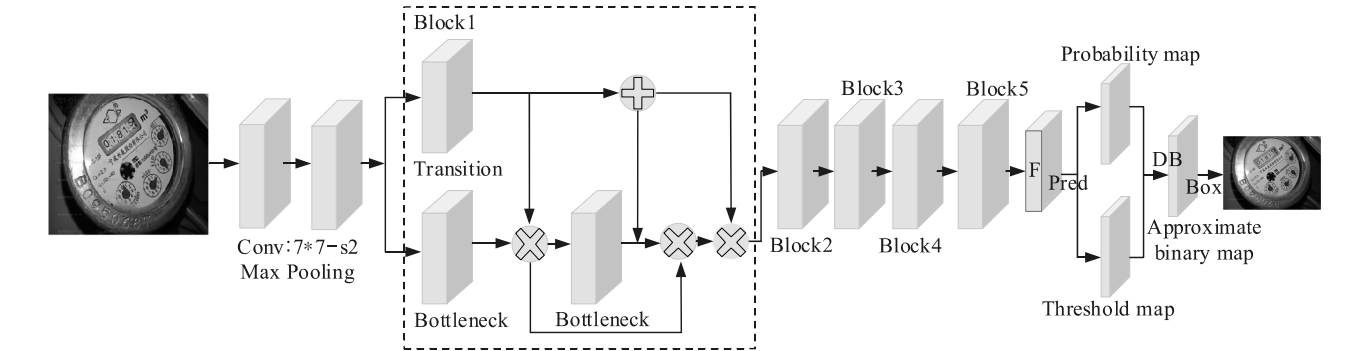


图 3 改进的 DBNet 网络模型结构

1.1.2 改进的 DBNet 训练结果分析

笔者选取 5 000 张不同类型的数字式水表图像数据，其中 80%作为训练集，20%作为验证集。为验证改进后 DBNet 网络模型的有效性，选取 R2CNN 网络、PSENet 网络、DBNet 网络、改进的 DBNet 网络 4 个算法的不同模型进行对比，即选取常用的模型性能评价指标：Precision、Recall、F1、FPS，

结果如表 1 所示。

表 1 不同网络模型性能对比

网络模型	Precision	Recall	F1	FPS
R2CNN	78.6	76.4	80.1	10.2
PSENet	86.1	76.2	81.7	11.7
DBNet+ResNet50	88.4	76.6	82.2	12.2
DPN+DBNet	91.2	76.0	85.6	15.4

由表 1 可知，在具有相同的数据和训练环境下，

改进的 DBNet 网络模型 Precision、F1、FPS 3 个指标都优于 R2CNN、PSENet、DBNet+ResNet50 网络模型, 虽然 Recall 指标低于其他对比模型几个百分点, 但是总体来说改进的 DBNet 模型比其他对比模型的检测效果更好、速度更快。

由图 4 可以看出: 基于改进的 DBNet 网络模型能够正确地检测出带有遮挡、弯曲和模糊的读数区域图像。



图 4 部分 DBNet 检测结果

综合表 1 和图 4 的检测结果可知: 基于改进的 DBNet 网络模型能够在消耗较少资源的情况下准确的检测不同类型的水表图像, 具有检测准确率高、性能好等优点。

1.2 读数区域的倾斜校正

为在读数识别过程中提高读数识别的准确率, 读数区域应该是水平的; 因此, 笔者使用透视变换原理^[9-10]对读数区域进行倾斜校正。透视变换需要原始图像固定的 4 个点坐标以及目标图像的 4 个点坐标, 通过原始图像 4 个点的坐标信息(tl , tr , br , bl), 计算出倾斜读数区域的最大宽度和最大高度, 假定目标读数区域的 4 个点的坐标为($[0, 0]$, $[maxWidth-1, 0]$, $[maxWidth-1, maxHeight-1]$, $[0, maxHeight-1]$), 计算出 2 个坐标之间的透视变换矩阵, 进行映射, 得到透视变换后的图像。

由于读数区域在水表图像中所处的位置不同, 透视变换后读数区域存在不同的水平方向; 因此, 需根据读数区域在水表图像所处的位置设定不同的阈值, 给定一个函数, 该函数的功能如下:

1) 比较读数区域的最大宽度和最大高度;

2) 若读数区域的最大高度大于最大宽度, 则判断读数区域左上角 x 的值是否小于预设的阈值(该阈值通过对多张数据集测试而来), 是, 则调用旋转函数将读数区域逆时针旋转 90° , 否, 则调用旋转

函数将读数区域顺时针旋转 90° ;

3) 若读数区域的最大高度小于最大宽度, 则判断读数区域左上角 y 的值是否大于预设的阈值, 是, 则调用旋转函数将读数区域顺时针旋转 180° , 否, 则说明读数区域是水平的不用旋转。

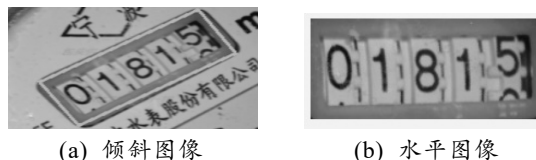


图 5 倾斜校正的过程

2 读数识别

2.1 数据集准备

实验数据来自于改进 DBNet 网络模型对原始水表图像进行检测后使用透视变换进行矫正的读数区域图像, 通过 labelme 标注工具对读数区域中若干个数字进行标注, 字轮部分全整字符按 0—9 这 10 个类别标注, 双半字符按 10—19 这 10 个类别标注, 部分读数区域标注如图 6 所示。最终共标注 2 633 张数据集, 其中清晰的图片有 1 945 张, 遮挡、模糊的图片有 397 张, 含有双半字符的图片有 291 张。

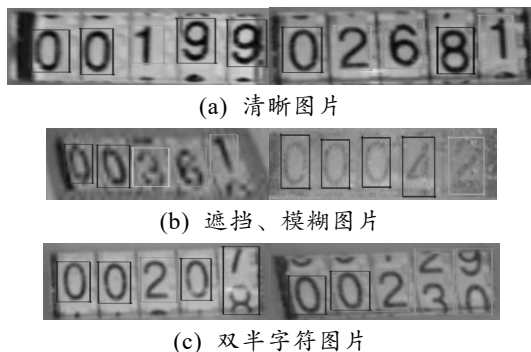


图 6 部分数据集标注

2.2 基于 YOLOv5 的读数识别

目标检测网络类型主要分为一阶段检测网络(one-stage)和二阶段检测网络(two-stage)。二阶段检测网络的识别准确率高于一阶段检测网络, 但是二阶段检测网络在生成候选区域时需要花费大量的时间, 导致效率低下、运行速度慢。一阶段检测网络无需生成候选框就可直接检测, 提高了训练速度。YOLO 作为一阶段检测网络的经典算法, 从 YOLOv1 发展至 YOLOv5, 每发展一个新的版本, 算法的性能都能随之提升, YOLOv5 综合了 YOLOv1 至 YOLOv4 的特性, 具有体积小、速度快等特点^[11-14]。为提高模型在应用平台上的检测速度, 笔者选取 YOLOv5s-v5.0 网络模型进行读数的

检测, 该网络结构由输入端、Backbone、Neck 和预测 4 部分组成, 网络结构如图 7 所示。

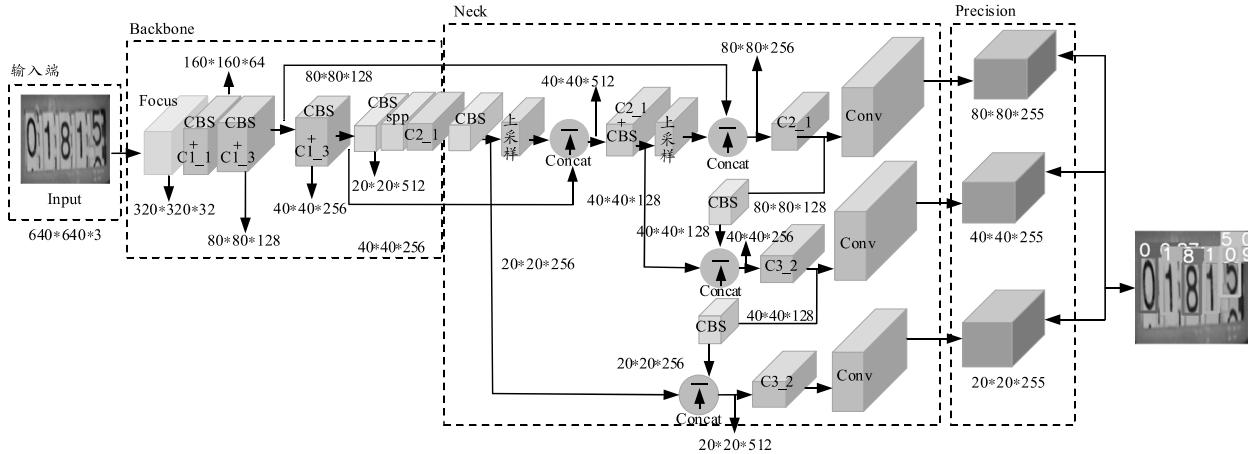


图 7 YOLOv5s 网络结构

具体流程为: 1) 在输入端中输入图片, 对输入图像进行图像预处理操作, 其中图像预处理包括马赛克数据增强, 提高对小目标的检测与识别; 添加自适应缩放处理, 将输入图像缩放为统一大小, 增强网络数据的处理能力。2) 通过 Focus 结构中的 4 个切片和一个 32 个卷积核的卷积, 将原来的 $640 \times 640 \times 3$ 图像转化为 $320 \times 320 \times 12$ 的特征图, 并使用 CSP^[15] 网络进行局部不同层之间融合, 获取到更多的特征信息。3) 通过 PANet 整合不同层之间的图像特征, 将不同的 CSP 网络顶部特征信息和输出的特征信息进行自上而下的特征融合, 然后将浅层特征自下向上进行特征融合。4) 采用加权非最大值抑制 (non-maximum suppression, NMS) 选择检测框, 得到检测结果。5) 利用训练好的权重检测读数区域中的数字, 其中部分检测结果如图 8 所示。

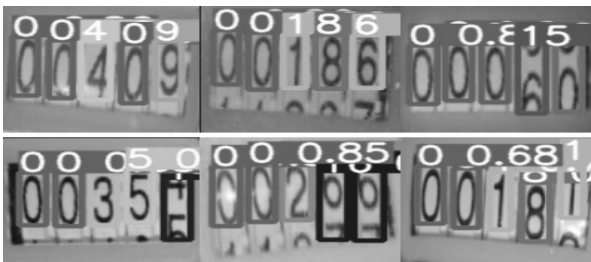


图 8 检测结果

图 8 中可以看出: 检测结果包括读数区域中每个字轮部分的检测类别和位置信息, 将检测结果存储在 R 中; 遍历 R 中 i 个字轮的类别信息 M_i , 位置信息 $P_i(x_{i1}, y_{i1}, x_{i2}, y_{i2})$, 通过计算每个节点 P_i -center(x_i, y_i) 的位置信息来获得相应的中心位置信息, 计算每个节点 P_i -center(x_i, y_i) 的公式如下:

$$\left. \begin{aligned} x_i &= (x_{i2} + x_{i1}) / 2 \\ y_i &= (y_{i2} + y_{i1}) / 2 \end{aligned} \right\} \quad (1)$$

式中 $i \in [1, n] \cap i \in Z$, 然后使用位置中心排序方法对每个字轮部分的位置信息进行排序, 并对每个字轮的类别信息 M_i 进行取余操作, 即为读数区域的读数结果。

2.3 YOLOv5s 网络模型训练结果分析

笔者选取 YOLOv5s 网络模型对读数区域中的数字进行训练, 共训练 2 633 张读数区域图像, 其中训练集、测试集和预测集的训练比例为 8:1:1, 训练后的 Precision、Recall 和 mAP 如图 9 所示。

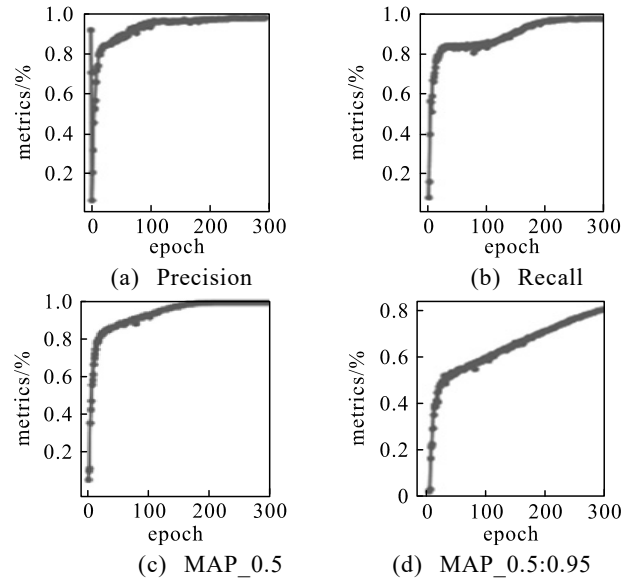


图 9 epoch 训练结果

由图 9 可以看出: 前 100 epochs 中, Precision、Recall 和 mAP 上升较快, 在训练 200 epochs 后 Precision、Recall 和 mAP@0.5 已经接近最优值, 直到训练 300 epochs 后 mAP@0.5:0.95 也接近最优值, 可达到准确识别数字的效果。模糊、遮挡和双半字符的检测结果如表 2 所示。

表 2 遮挡、模糊和双半字符检测结果

数据类型	图纸张数	错误识别张数	正确识别张数
清晰	1 945	0	1 945
遮挡、模糊	397	10	387
双半字符	291	4	287

由表 2 可以看出：该模型对读数区域中带有遮挡、模糊以及双半字符的图像均能够正确地检测出来，准确率达到 99.1%。

3 水表识别应用系统

水表识别应用系统基于水务公司对城镇居民用水量的管理角度以及水务公司客户查看用水量的角度，开发了服务器端和手机端的水表识别应用系统。其中服务器端主要用于水务公司对水表数据进行管理，更为重要的是手机端若存在检测不出来的图像，对服务器端提出报警，服务器端对存在可疑的水表图像进行处理并将结果反馈给手机端；手机端则直接面向水务公司管理的客户，客户通过拍照上传水表图像，手机端反馈用水量以及需缴纳的费用给客户。

服务器端水表识别界面如图 10 所示。该界面为服务器端的水表读数识别模块的结果展示，具体流程为：1) 通过上传水表图像，将图像输入到读数区域定位模块，对图像的读数区域进行检测；2) 将读数区域的图像输入到读数识别模块，对读数区域中的数字进行检测；3) 将水表图像的检测结果进行显示。

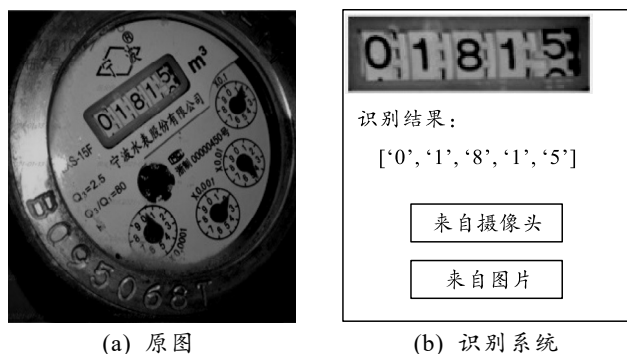


图 10 服务器端水表识别界面

手机端水表识别应用系统界面如图 11 所示，界面由 2 个按钮组成。客户可以选择拍照按钮或者相册已有水表图像进行上传，系统则对上传的图片进行水表读数识别。

由图 10 和 11 可以看出：水表识别应用系统搭建笔者设计的读数区域检测算法和读数识别算法，能够对输入的水表图像准确的进行检测和识别，但也可能存在少量检测错误的水表图像；因此，需通

过水务公司的工作人员对出现错误的水表图像进入水表识别应用系统后台进行人工的抄表，并对抄表结果进行记录和反馈。

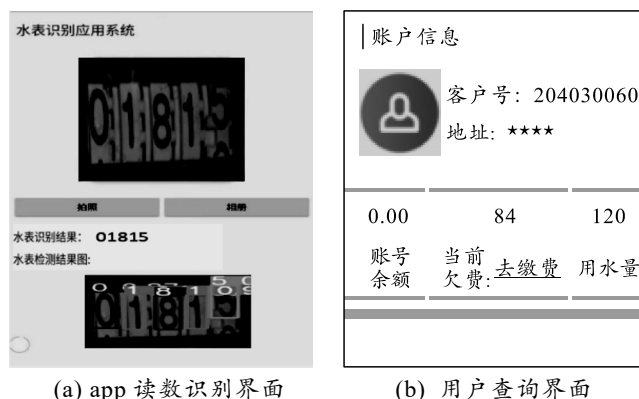


图 11 手机端水表识别界面

4 结论

为方便水务公司对客户用水量的管理，笔者针对不同环境下拍摄的多角度水表图像，提出一种两阶段双模型的水表读数识别。实验结果表明：在读数区域定位方面改进 DBNet 网络模型不仅能够对污染严重的水表图像中的读数区域进行检测，而且相比其他的网络模型性能更好、速度更快；在读数识别方面，YOLOv5s 网络模型对读数区域中数字识别的准确率达到 99.1%；再搭建水表识别应用系统，实现了水表的自动读数功能。使用笔者设计的水表识别应用系统弥补了传统人工抄表方式费时、费力、效率低的缺点，提高了水务公司在居民用水方面的管理效率和经济成本。

参考文献：

- [1] 汤鹏, 刘毅, 魏宏光, 等. 基于 Mask-RCNN 海上升压站数字式仪表读数的自动识别算法[J]. 红外与激光工程, 2021, 50(S2): 163-170.
- [2] 侯卓成, 欧阳华, 胡鑫, 等. 基于改进的 YOLOv4 彩色数字仪表读数识别方法[J]. 电子测量技术, 2022, 45(6): 124-129.
- [3] 华泽玺, 施会斌, 罗彦, 等. 基于轻量级 YOLO-v4 模型的变电站数字仪表检测识别[J]. 西南交通大学学报, 2022, 12(8): 1-11.
- [4] 郭兰英, 韩睿之, 程鑫. 基于可变形卷积神经网络的数字仪表识别方法[J]. 计算机科学, 2020, 47(10): 187-193.
- [5] 周曼, 刘志勇, 鲁乾鹏, 等. 基于 OCR 的数字仪表自动识别在工业现场中的应用[J]. 仪器仪表用户, 2021, 28(1): 22-24.