

## 基于Mask-RCNN海上升压站数字式仪表读数的自动识别算法

汤鹏 刘毅 魏宏光 董秀芬 严国斌 张迎宾 袁亚君 王增光 范亚南 马鹏阁

### Automatic recognition algorithm of digital instrument reading in offshore booster station based on Mask-RCNN

Tang Peng, Liu Yi, Wei Hongguang, Dong Xiufen, Yan Guobin, Zhang Yingbin, Yuan Yajun, Wang Zengguang, Fan Yanan, Ma Pengge

在线阅读 View online: <https://doi.org/10.3788/IRLA20211057>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 海上升压站自动巡检机器人颜色检测算法研究

Research on color detection algorithm of automatic patrol robot in offshore booster station

红外与激光工程. 2020, 49(S2): 20200589–20200589 <https://doi.org/10.3788/IRLA20200589>

#### 基于特征选择YOLOv3网络的红外图像绝缘子检测方法

Insulator detection method based on feature selection YOLOv3 network

红外与激光工程. 2020, 49(S2): 20200401–20200401 <https://doi.org/10.3788/IRLA20200401>

#### 海洋激光雷达图像处理提取海水深度的方法

Extracting sea water depth by image processing of ocean lidar

红外与激光工程. 2021, 50(6): 20211034–1–20211034–8 <https://doi.org/10.3788/IRLA20211034>

#### 图像处理在光纤光斑微位移传感中的应用

Application of image processing in micro-displacement sensing of fiber speckle

红外与激光工程. 2018, 47(10): 1022004–1022004(7) <https://doi.org/10.3788/IRLA201847.1022004>

#### 红外偏振感知与智能处理

Infrared polarization perception and intelligent processing

红外与激光工程. 2018, 47(11): 1102001–1102001(7) <https://doi.org/10.3788/IRLA201847.1102001>

#### 图像能量与对比度的霾检测算法

Haze detection algorithm based on image energy and contrast

红外与激光工程. 2018, 47(8): 826001–0826001(6) <https://doi.org/10.3788/IRLA201847.0826001>

## 基于 Mask-RCNN 海上升压站数字式仪表读数的自动识别算法

汤 鹏<sup>1,2</sup>, 刘 毅<sup>3</sup>, 魏宏光<sup>2</sup>, 董秀芬<sup>1</sup>, 严国斌<sup>4</sup>, 张迎宾<sup>4</sup>, 袁亚君<sup>4</sup>, 王增光<sup>3</sup>, 范亚南<sup>3</sup>, 马鹏阁<sup>2</sup>

- (1. 中国长江三峡集团有限公司, 北京 100038;
2. 郑州航空工业管理学院 智能工程学院, 河南 郑州 450015;
3. 中国航空工业集团公司洛阳电光设备研究所, 河南 洛阳 471000;
4. 三峡新能源海上风电运维江苏有限公司, 江苏 盐城 224008)

**摘要:** 海上升压站采用挂轨机器人开展巡检作业, 利用机器视觉手段自动识别数字式仪表读数, 替代人工记录。提出了一种基于 Mask-RCNN 深度学习方法的数字仪表读数自动识别算法。将不同类型的数字仪表原始图像制作成数据集, 利用深度学习算法进行训练, 根据损失函数变化曲线对算法进行参数优化得到训练后的模型, 再进行数字仪表图像的识别分析。采用灰度世界算法和霍夫变换等算法进行图像预处理, 可有效改善数字识别的准确度。最后, 实验对比了 YOLOv3 和 Mask-RCNN 深度学习算法的识别性能, 结果表明前者具有较高的检测速度, 后者具有更高的准确率。后者的识别率为 99.52%, 满足海上升压站远程监控对数字仪表读数正确率高的要求。

**关键词:** 图像处理; 数字式仪表识别; Mask-RCNN; YOLOv3

**中图分类号:** TP391.41; TP242; P75 **文献标志码:** A **DOI:** 10.3788/IRLA20211057

## Automatic recognition algorithm of digital instrument reading in offshore booster station based on Mask-RCNN

Tang Peng<sup>1,2</sup>, Liu Yi<sup>3</sup>, Wei Hongguang<sup>2</sup>, Dong Xiufen<sup>1</sup>, Yan Guobin<sup>4</sup>, Zhang Yingbin<sup>4</sup>, Yuan Yajun<sup>4</sup>, Wang Zengguang<sup>3</sup>, Fan Yanan<sup>3</sup>, Ma Pengge<sup>2</sup>

- (1. China Three Gorges Corporation, Beijing 100038, China;
2. School of Intelligent Engineering, Zhengzhou Institute of Aeronautics Industry Management, Zhengzhou 450015, China;
3. Luoyang Institute of Electro-Optical Equipment, Aviation Industry Corporation of China, Luoyang 471000, China;
4. Three Gorges New Energy Offshore Wind Power Operation and Maintenance Jiangsu Co., Ltd, Yancheng 224008, China)

**Abstract:** The offshore booster station adopts the rail hanging robot to carry out patrol inspection, and the machine vision method is used to automatically identify the digital instrument reading instead of manual recording. An automatic recognition algorithm of digital instrument reading based on Mask-RCNN deep learning method was presented. The original images of different types of digital instruments were made into data sets, trained by deep learning algorithm, the parameters of the algorithm were optimized according to the change curve of loss function, the trained model was obtained, and then the digital instrument images were recognized and analyzed. The gray world algorithm and Hough transform were used for image preprocessing, which can effectively improve the accuracy of digital recognition. Finally, the recognition performance of YOLOv3 and Mask-RCNN deep learning algorithm was compared in the experiment. The results show that the former has

收稿日期: 2021-05-10; 修订日期: 2021-06-23

基金项目: 中国长江三峡集团有限公司科研资助项目 (202003036)

作者简介: 汤鹏, 男, 工程师, 博士, 主要从事机器人导航算法开发、电驱动系统设计等方面的研究。

higher detection speed and the latter has higher accuracy. The recognition rate of the latter is 99.52%, it meets the requirement that remote monitoring of offshore booster station requires high accuracy of digital instrument reading.

**Key words:** image processing; digital instrument recognition; Mask-RCNN; YOLOv3

## 0 引言

海上升压站由于其地理位置的特殊性,使其设备维护、常规检查等任务难度加大且耗时耗力。由于海上升压站结构固定,空间紧凑,较适合选用挂轨机器人进行巡检作业。挂轨机器人搭载具有上下升降功能的全方位云台,安装可见光相机实现对室内各设备的近距离全方位巡检。升压站中使用的工业仪表大部分为数字式仪表。数字式仪表具有读数准确方便、测量速度快、能提供数字信号输出、方便进行目视观察和数字记录等优点,在升压站的机房中也被大量应用。然而,海上升压站距离海岸远,工作环境不适合人工采集,亟需开发自动读数识别算法。

海上升压站的数字式仪表数字识别是自动巡检中的重要任务。数字式仪表识别的方法经过多年的发展,在这一领域也取得了一些进展。国内较早开始数字仪表识别的是段会川等<sup>[1]</sup>,提出了一种基于模糊理论的仪表数字识别方法,构建了一种基于模糊算法的数字识别器,能够快速进行数字的识别,但对采集图片的质量要求较大,对于倾斜图像不具备较好的分割效果,进而可能会造成数字的残缺,影响识别效果。郭爽<sup>[2]</sup>、卢卫娜等<sup>[3]</sup>都是用模板匹配的方法进行数字识别,针对参考文献<sup>[2]</sup>中传统的模板匹配方法对因干扰稍有变形、位移、旋转的图像难以判别的问题,参考文献<sup>[3]</sup>中提出了一种改进的标准模板匹配方法,对每类数字的各种干扰进行细分,制作多个模板进行匹配,在一定程度上降低了因干扰而带来的识别误差,但制作的模板数量大大增加,使方法复杂度增加。李素萍<sup>[4]</sup>提出了一种基于图像处理技术的数字仪表识别方法,利用模式识别的方法对数字进行识别,对于比较清晰、完整、竖直的数字具有很好的识别率。近年来,深度学习的神经网络发展迅速,在数字式仪表识别领域引入了卷积神经网络(Convolutional Neural Network, CNN)<sup>[5]</sup>,它对于单个数字的识别效果较好,但对于干扰较多的数字式仪表的数字识别,准确率较低。目前应用比较多的目标检测算法包

括 RCNN(Regions with CNN features)<sup>[6]</sup>系列、YOLO(You Only Look Once)<sup>[7]</sup>系列、SSD(Single Shot Multibox Detector)<sup>[8]</sup>等,它们在一些大目标、轻量化的应用场景中前景广阔,具有较高的精确度,但对于一些尺度较小的目标检测有所缺陷。

海上升压站的内部结构较适合使用挂轨机器人,机器人可对数字式仪表进行定位拍摄,在一定程度上减少外在环境的干扰,同时数字式仪表图像数字像素信息少,数字间的粘连大而造成分割较难。针对以上的问题,提出了基于 Mask-RCNN 网络的数字仪表识别方法,Mask-RCNN 网络在 Faster-RCNN 的基础上增加了一个 FCN 语义分割网络分支,用于对目标进行像素级的语义分割。用 ROI Align 层替换 ROI Pooling 层,取消了量化操作,避免了量化引入的误差,提升了精度。同时使用基于残差网络 ResNet101 和特征金字塔网络 (Feature Pyramid Network, FPN)<sup>[9]</sup>共同组成的特征提取网络替代 CNN 特征提取网络,优化了目标检测中的多尺度问题,提高了小物体、像素信息少等目标的检测性能。

## 1 数字仪表识别方法

### 1.1 模板匹配法

基于模板匹配的数字识别方法<sup>[3]</sup>首先是对采集的图像进行直方图增强、滤除噪声等预处理,然后构建改进的仪表数字模板,进行数字的识别。整个过程较为复杂,需要构建多个模板,同时若采集到的数字出现变形、位移、旋转等情况,会造成数字难以被识别,可靠性较低。

### 1.2 穿线法

由于仪表数字大多是由七段数码管组合而成的,只有横竖的布置方式,可以用基于穿线法进行仪表数字的特征识别<sup>[10]</sup>。该方法通过图像预处理、数字的定位进行分割,再将分割的单个数字使用穿线识别法进行数字的识别。该方法较其他方法原理简单,但对数字分割要求过高,数字在比较规整的条件下才会有较

好的识别效果,且不能识别小数点。

### 1.3 基于机器学习的数字识别

机器学习是模式识别的重要方法。将采集到的图片经过倾斜校正、形态学处理、基于连通域的方法来进行数字的定位与分割,最后将分割后的单个字符使用 LSSVM 算法识别仪表数字<sup>[11]</sup>。机器学习的方法相比前两种方法有着更好的适应性和识别率。

### 1.4 基于深度学习的数字识别

近年来,计算机视觉发展迅速,基于深度学习的目标检测方法逐渐成为主流方法。目前常用的目标检测深度学习算法有 YOLO 系列和 R-CNN 系列算法。YOLO 算法在大目标、轻量化的场景中具有较好的检测效果,同时具有较快的检测速度,但对小目标的检测效果不是很好。R-CNN 网络在检测的精度方面具有较好的效果,Mask-RCNN 网络采用特征金字塔网络进行多维度特征提取,同时增加了一个 FCN 语义分割网络,在完成目标定位和分类的同时,实现对目标轮廓像素级的语义分割,大大提高了目标检测的精度,但同时减弱了目标检测的速度。文中主要针对深度学习用于数字识别开展应用研究。

## 2 基于深度学习的数字仪表识别

### 2.1 图像预处理

由于数字式仪表屏幕大多数是由液晶制作成的,在阳光、灯光等其他光照条件的照射下,屏幕易发生反光,造成数字的过亮或者过暗,或者数字模糊、亮度不均衡等问题。同时挂轨机器人对数字式仪表的图像采集中,由于拍摄角度不同,采集到的图像可能会存在一定的倾斜。亮度不均衡、倾斜等问题都会影响数字式仪表数值的识别,为了减弱这种影响,提高识别的精度,需要对采集到的图像进行预处理。

#### 2.1.1 屏幕区域矫正

实验使用 Gray World 算法<sup>[12]</sup>来减弱因光照等因素引起的数显区域亮度不均衡的情况。人的视觉系统具有颜色恒常性,能从变化的光照环境和成像条件下获取物体表面颜色的不变特性,但成像设备不具有这样的调节功能,不同的光照环境会导致采集的图像颜色与真实颜色存在一定程度的偏差,需要选择合适的颜色校正算法,消除光照环境对颜色显现的影响。

Gray World 算法以灰度世界假设为基础,该假设

认为:对于一幅有着大量色彩变化的图像,其  $R, G, B$  3 个分量的平均值趋于同一灰度值。从物理意义上讲,灰色世界假设自然界景物对于光线的平均反射的均值在总体上是定值,这个定值近似地为“灰色”。颜色平衡算法将这一假设强制应用于待处理图像,可以从图像中消除环境光的影响,获得原始场景图像。该算法的具体步骤如下。

(1) 求反射均值  $\overline{Gray}$

$$\overline{Gray} = \frac{\overline{R} + \overline{G} + \overline{B}}{3} \quad (1)$$

式中:  $\overline{R}, \overline{G}, \overline{B}$  为彩色图像在  $R, G, B$  3 个色彩通道上的亮度平均值;

(2) 通过步骤 (1) 计算得到的 3 个色彩通道亮度的平均值,计算  $R, G, B$  3 个通道的增益系数  $k_r, k_g, k_b$ , 计算公式为:  $K_r = \frac{\overline{Gray}}{\overline{R}}; K_g = \frac{\overline{Gray}}{\overline{G}}; K_b = \frac{\overline{Gray}}{\overline{B}};$

(3) 根据 Von-Kries 对角模型,对于图像中的每个像素  $C$ ,调整其  $R, G, B$  分量为:

$$\begin{cases} C(R') = C(R) \times k_r \\ C(G') = C(G) \times k_g \\ C(B') = C(B) \times k_b \end{cases} \quad (2)$$

通过 GrayWorld 算法得到亮度均衡后的图像,使数显区域的数字更加清晰。在一定程度上减弱了因不同光照条件下对图像数显区域成像效果的影响,提高了识别的准确度。GrayWorld 算法进行亮度均衡前后的图像如图 1 所示。

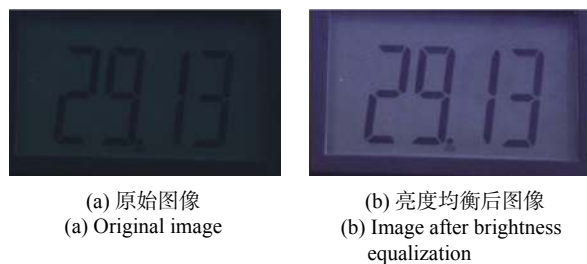


图 1 GrayWorld 算法亮度均衡前后图像

Fig.1 Images before and after brightness equalization with GrayWorld algorithm

#### 2.1.2 图像倾斜校正

实验采用 Canny 算子进行数显区域的边缘检测,采用 Hough 变换获取数显区域边框的倾斜角,然后再采用旋转函数对图像进行校正<sup>[13]</sup>。

采集到的数字式仪表图像中,数显区域的上下边

界存在两条平行直线,将平行线与边框的夹角设定为倾斜角,即为采集图像的倾斜角。因为数字式仪表图像中数显区域的边框是最长的一条直线,所以通过累加器计算 Hough 变换得到的最大值对应的 $(\rho, \theta)$ ,由此得到采集图像所对应的倾斜角度。倾斜图像转换为端正图像的效果如图 2 所示。

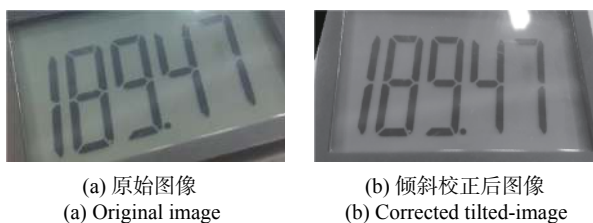


图 2 校正倾斜图像的实验效果

Fig.2 Corrected tilted-image experiment effect

### 2.2 基于 YOLOv3 网络的数字仪表识别算法

YOLOv3 检测算法<sup>[14]</sup>于 2018 年被首次提出,将一个图像输入网络,就可以得到图像中目标的位置和它们的类别以及相应的置信度。它的训练和检测、特征提取、分类回归都是在一个网络中完成的,真正实现端到端的目标识别。由于海上升压站仪表位置固定,挂轨机器人根据位置坐标对数显区域进行图像采集,输入到 YOLOv3 网络进行数字识别。数字识别流程如图 3 所示。

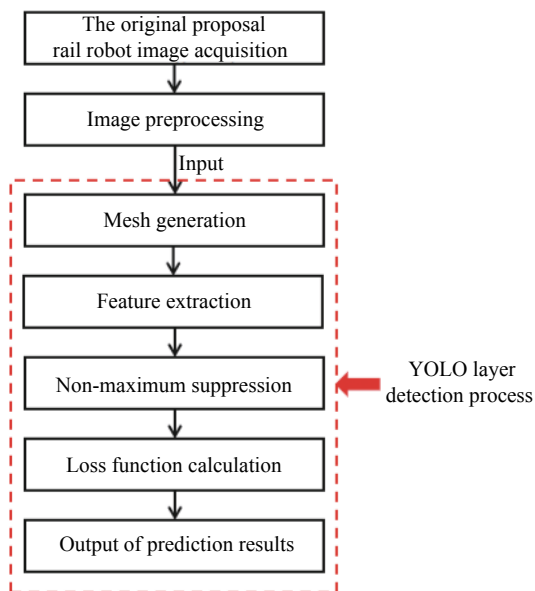


图 3 YOLOv3 数字识别流程

Fig.3 YOLOv3 digital identification process

数字式仪表识别步骤如下:

(1) 预处理。通过对图像进行预处理减弱因外界环境对数字识别的误差。

(2) 输入一张任意大小图片,将图像的分辨率缩放至  $a \times b$ ,通道数为 3 的 RGB 图像,将生成的新图像作为网络输入 input。

(3) 卷积神经网络特征提取。YOLOv3 的卷积神经网络(CNN)把图像划分成  $S \times S$  个网格(grid cell)。YOLOv3 进行多尺度预测,输出有 3 层的 feature maps,每层 feature maps 有  $S \times S$  个网格,分别为  $13 \times 13$ 、 $26 \times 26$  和  $52 \times 52$ ,每个网格负责去检测那些中心点落在该格子内的目标。特征图的输出维度为  $S \times S \times [3 \times (5 + 11)]$ , $S \times S$  为输出特征图点数,一共 3 个 anchor 框,每个框有 4 维预测框数值 $(t_x, t_y, t_w, t_h)$ ,1 维预测框置信度,11 维物体类别数。

预测框的置信度 Confidence 用公式表示为:

$$\text{Confidence} = P_r(\text{Object}) \times IOU_{pred}^{truth} \quad (3)$$

$$IOU_{pred}^{truth} = \frac{\text{area}(t) \cap \text{area}(p)}{\text{area}(t) \cup \text{area}(p)} \quad (4)$$

式中:  $P_r$  为目标 bounding box 内存在目标的概率;  $IOU_{pred}^{truth}$  为真实框和预测框面积的交并比;  $\text{area}(t)$  为真实框的面积;  $\text{area}(p)$  为预测框的面积。

(4) 非极大值抑制。YOLOv3 对于每个单元格有 3 个预测框,通过非极大抑制方法,设置一个 Score 的阈值,低于该阈值的预测框置信度设置为 0,遍历每一个对象类别,选择得分最高的作为输出,与该输出重叠的去掉,不断重复这一过程直到所有预测框处理完。在每个网格中,对象  $C_i$  位于第  $j$  个 bounding box 的得分为:

$$\text{Score}_{ij} = P_r(C_i | \text{Object}) \times \text{Confidence}_{ej} \quad (5)$$

(5) 损失函数<sup>[15]</sup>。YOLO 网络的损失函数由 3 部分构成,分别是预测框和真实框的尺寸误差、预测框和真实框的置信度误差、分类误差。

预测框和真实框的尺寸误差 loss1 用公式表示为:

$$\text{loss1} = \lambda_{coord} \sum_{i=0}^{S \times S} \sum_{j=0}^B I_{ij}^{obj} (2 - w_i \times h_i) \times \left[ (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2 \right] \quad (6)$$

预测框和真实框的置信度误差 loss2 用公式表示为:

$$\text{loss}_2 = \sum_{i=0}^{S \times S} \sum_{j=0}^B I_{ij}^{obj} (c_i - \hat{c}_i)^2 + \lambda_{noobj} \sum_{i=0}^{S \times S} \sum_{j=0}^B I_{ij}^{noobj} (c_i - \hat{c}_j)^2 \quad (7)$$

分类误差 loss3 用公式表示为:

$$\text{loss}_3 = \sum_{i=0}^{S \times S} I_{ij}^{obj} \sum_{C \in \text{classes}} (p_i(C) - \hat{p}_i(C))^2 \quad (8)$$

YOLO 网络总的损失函数为:  $\text{loss} = \text{loss}_1 + \text{loss}_2 + \text{loss}_3$ 。

### 2.3 基于 Mask-RCNN 网络的数字仪表识别算法

Mask-RCNN 框架是由 Facebook 人工智能研究小组提出的端到端 CNN, 是基于 Faster-RCNN 的网络结构改进而来的, 其框架结构简单易操作, 且识别精度指标十分优越, 适合数字式仪表图像识别这类像素特

征较少、读数精度需求高的任务。其网络结构如下<sup>[16]</sup>:

(1) 图片采集, 由于 Mask-RCNN 对硬件性能需求较高, 需要将挂归机器人采集的图片通过局域网传入终端服务器进行处理。

(2) 图片的预处理, 同 2.2 节处理, 这里不再赘述。

(3) 输入预处理后的图片, 使用 FPN 网络进行特征提取。网络结构如图 4 所示。基本特征提取由 ResNet(残差网络) 完成, ResNet 通过 4 个 block 进行特征提取、输出及保存, 构成 list 1, 对 list1 中 c5 进行上采样, 生成 list2 [p2,p3,p4,p5,p6]。对生成的 feature map 中每一点设定预定 ROI, 从而获得多个候选 ROI。

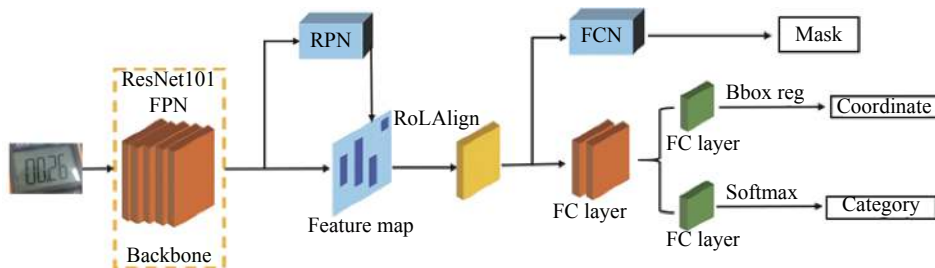


图 4 Mask-RCNN 结构

Fig.4 Mask-RCNN structure

(4) 将候选 ROI 送入 RPN 网络二值分类, 通过 RPN 生成目标矩阵块, 过滤掉部分 ROI。将 list2 依次输入并输出, 得到不同尺度 proposals box(候选区域)。基于 anchor 预测中心偏差及长宽偏差, 得到与真实窗口相近的回归窗口, 如图 5 所示。

边框回归公式如下:

$$(\Delta x, \Delta y), \Delta x = p_w d_x(p), \Delta y = p_h d_y(p) \quad (9)$$

$$\begin{aligned} \hat{G}_x &= p_w d_x(p) + p_x \\ \hat{G}_y &= p_h d_y(p) + p_y \end{aligned} \quad (10)$$

(5) 通过 ROI Align 层, 将相应区域池化为固定尺寸的特征图, 从而进行后续的分类和包围框回归操作。ROI Align 是一种区域特征聚集方式, 工作原理如图 6 所示。它取消了量化操作, 使用双线性内插的方法<sup>[17]</sup> 获得坐标为浮点数的像素点上的图像数值, 最终将 ROI 再通过 bbox 进行边框回归以及 softmax 分类器进行分类。

双线性内插的计算公式为:

$$\begin{aligned} f(R_1) &\approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}) \\ \text{where } R_1 &= (x, y_1) \\ f(R_2) &\approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}) \\ \text{where } R_2 &= (x, y_2) \end{aligned} \quad (11)$$

(6) Mask 层也是 Mask-RCNN 网络中特有部分,

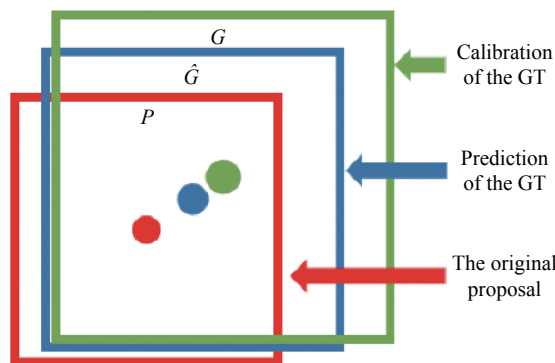


图 5 回归窗口

Fig.5 Return to the window

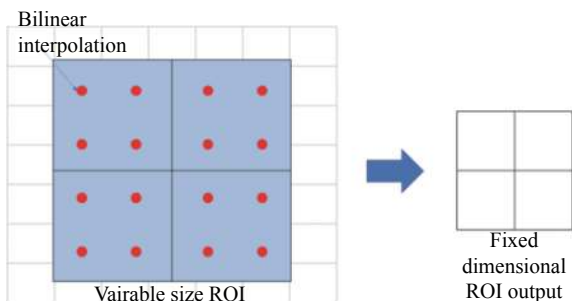


图 6 Feature map 图示

Fig.6 Feature map graphic

在 Mask 分支上, 添加 FCN 网络, 对每个 RoI 的分割输出维数为  $K \times m \times m$  (其中:  $m$  表示 RoI Align 特征图的大小), 即  $K$  个类别的  $m \times m$  的二值 Mask; 保持  $m \times m$  的空间布局, pixel-to-pixel 操作需要保证 RoI 特征映射到原图的对齐性, 这也是使用 RoIAlign 解决对齐问题的原因, 能够减少像素级别对齐的误差, 这种方式能够有效提高实例分割的效果。Mask-RCNN 的损失函数由 3 项损失函数组成, 包括分类误差、回归误差和分割误差, 总的损失函数表达式为:

$$\text{loss} = \text{loss}_{cls} + \text{loss}_{reg} + \text{loss}_{mask} \quad (12)$$

### 3 实验与结果分析

#### 3.1 实验环境

实验在内存为 32 G 的 Windows10 专业版 64 位操作系统中进行, CPU 处理器为 Intel Xeon E5-2687Wv4, GPU 处理器为 NVIDIA GeForce GTX Titan-X, 选取 TensorFlow 作为学习框架。

#### 3.2 实验数据

数字式仪表数据: 此次实验使用的数据集均为实测数据, 该数据集包括: 0~9 和小数点共 11 类, 每类各 300 张做训练集和 30 张做测试集。数据集中包括不同角度、不同距离、不同光照强度的数据。数据集如图 7 所示。

YOLOv3 的数据集使用 labeling 软件进行标注, 标注标签为 0~9 和小数点共 11 类, 数据集格式为 VOC 格式。训练 Mask-RCNN 需要原始图像, 还要与之相对应的掩模信息, 因此先使用 Labelme 工具进行数据标注和掩模制作, 数据标注方式为手动标注, 数据集格式为 COCO 格式。实验数据标签为 0~9 和小数点共 11 类, 相较于传统数字式仪表识别算法,



图 7 部分数据集

Fig.7 Partial data set

Mask-RCNN 算法是针对每个像素的语义分割, 大大提高了数字识别的准确度。

#### 3.3 实验结果分析

使用 Mask-RCNN 网络在数据集上训练的损失函数变化曲线如图 8 所示。图 8(a) 为网络的总体损失变化曲线; 图 8(b) 为边框回归的损失变化曲线; 图 8(c) 为掩模的损失变化曲线; 图 8(d) 为分类损失变化曲线。图中横坐标表示网络权重更新的次数。

表 1 为实验数据集通过两种不同方法的识别率和耗时。YOLOV3 是一种基于深度神经网络的对象识别和定位算法, 其最大的特点是运行速度快, 可以用于实时系统, 但对小目标的检测精度不是很好。Mask-RCNN 在 Faster-RCNN 的基础上增加了一个分支用于分割任务, 对于每个 Proposal Box 都要使用

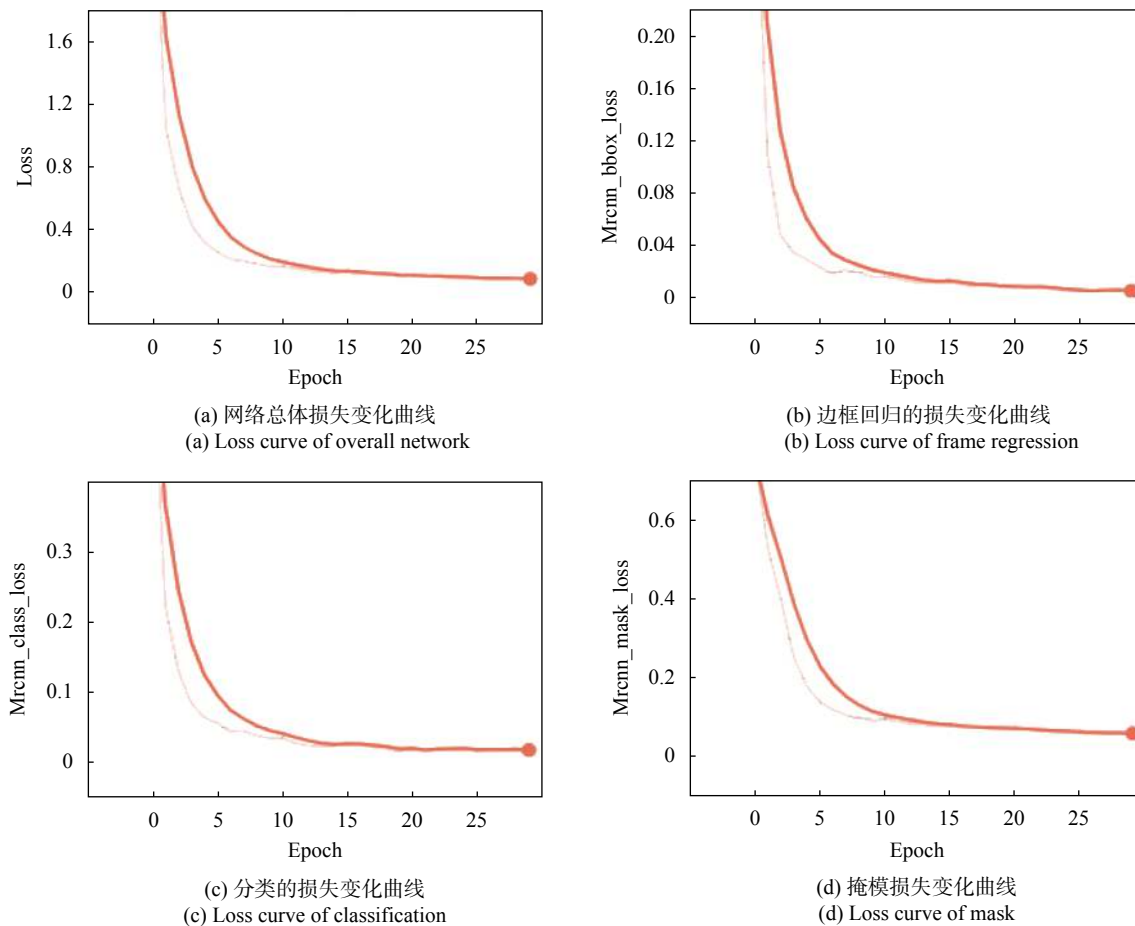


图 8 损失变化曲线

Fig.8 Loss function curve

表 1 不同识别方法在数据集上的比较

Tab.1 Comparison of different recognition methods on data sets

Model	Test set/frame	Accuracy	Time consuming/ms
YOLOv3	100	99.03%	20.2
Mask-RCNN	100	99.52%	212

FCN 进行语义分割, 分割任务与定位、分类任务是同时进行的, 同时引用 ROI Align 代替 ROI Pooling 提高了其检测的精度。此次实验采用这两种方法对数字仪表数字集进行识别, 实验结果如图 9 所示。



图 9 实验结果图

Fig.9 Experimental result diagram

#### 4 结论

针对海上升压站距离海岸远, 工作环境复杂, 人工采集数字式仪表数值较困难的问题, 文中分别使用 YOLOv3 算法和 Mask-RCNN 算法对数字式仪表进行数字识别。先将采集到的图片进行预处理, 对图像进行校正, 然后输入到检测网络中进行特征提取、分类和回归, 同时将置信度标注要图像上。经过对比分析发现, YOLOv3 算法进行数字识别具有耗时短的优势, 可用于需要实时检测的工程中, 而 Mask-RCNN 算法进行数字识别具有精度高的优势, 但检测速度较



慢,可用于对精度有较高要求的工程中。

### 参考文献:

- [1] Duan Huichuan, Zhang Haibo, Zhang Shuguang, et al. Research on instrument digital recognition based on fuzzy theory [J]. *Instrument Technique and Sensor*, 2004(4): 37-39. (in Chinese)
- [2] Guo Shuang. Research on automatic identification method of nixie tube digital instrument [J]. *Communications Technology*, 2004, 45(4): 37-39. (in Chinese)
- [3] Lu Weina, Liu Changrong, Zheng Yucai, et al. A digital instrument character recognition method based on template matching [J]. *Modern Computer*, 2008(3): 70-72, 86. (in Chinese)
- [4] Li Suping. Digital instrument recognition technology based on image processing [J]. *Mechanical and Electrical Engineering*, 2013, 19(6): 84-86, 90. (in Chinese)
- [5] Akula A, Singh A, Ghosh R, et al. Target recognition in infrared imagery using convolutional neural network[C]//Proceedings of International Conference on Computer Vision and Image Processing, 2017.
- [6] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//CVPR. IEEE, 2014.
- [7] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Computer Vision & Pattern Recognition. IEEE, 2016.
- [8] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot MultiBox detector[C]//European Conference on Computer Vision. Springer International Publishing, 2016.
- [9] Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.
- [10] Zhuo Lei, Zhou Lv, Yang Lihong. Numeral recognition of calculator based on threading method [J]. *Software Engineering*, 2018, 21(12): 1-3. (in Chinese)
- [11] Lin Jianping, Liao Yipeng. Automatic recognition of digital instrument reading based on OpenCV and LSSVM [J]. *Microcomputer & its Applications*, 2017, 36(2): 37-40. (in Chinese)
- [12] Guo Lanying, Han Ruizhi, Cheng Xin. Digital instrument recognition method based on deformable convolutional neural network [J]. *Computer Science*, 2020, 47(10): 187-193. (in Chinese)
- [13] Liu Jing. Research on automatic character recognition of digital instrument based on image processing [J]. *Information Technology*, 2020, 44(4): 84-87, 91. (in Chinese)
- [14] Redmon J, Farhadi A. YOLOv3: An Incremental improvement [J]. *arXiv e-prints*, 2018: 1804.02767.
- [15] Gong An, Zhang Yang, Tang Yonghong. Identification method of electric energy representation based on yolov3 network [J]. *Computer Systems & Applications*, 2020, 29(1): 196-202. (in Chinese)
- [16] He Peilin. Research and application of reading recognition algorithm for industrial instruments based on deep learning[D]. Chengdu: University of Electronic Science and Technology of China, 2020. (in Chinese)
- [17] Wang Sen, Yang Kejian. Research and implementation of image scaling algorithm based on bilinear interpolation [J]. *Techniques of Automation and Applications*, 2008(7): 44-45, 35. (in Chinese)