

基于边界约束图像融合的光学字符识别算法研究

张驯 赵金雄 白万荣 赵红

Research on optical character recognition algorithm based on boundary constrained image fusion

Zhang Xun, Zhao Jinxiong, Bai Wanrong, Zhao Hong

在线阅读 View online: <https://doi.org/10.3788/IRLA20220102>

您可能感兴趣的其他文章

Articles you may be interested in

图像融合在偏振关联成像中的应用

Application of image fusion in polarization correlated imaging

红外与激光工程. 2018, 47(12): 1226002 <https://doi.org/10.3788/IRLA201847.1226002>

红外与可见光图像融合技术的研究进展

Research progress of infrared and visible image fusion technology

红外与激光工程. 2021, 50(9): 20200467 <https://doi.org/10.3788/IRLA20200467>

基于FDST和双通道PCNN的红外与可见光图像融合

Infrared and visible image fusion based on FDST and dual-channel PCNN

红外与激光工程. 2019, 48(2): 204001 <https://doi.org/10.3788/IRLA201948.0204001>

制冷型中波红外偏振成像光学系统设计

Design of cooled medium-wave infrared polarization imaging optical system

红外与激光工程. 2021, 50(2): 20200208 <https://doi.org/10.3788/IRLA20200208>

图像融合在空间目标三维重建中的应用

Application of image fusion in 3D reconstruction of space target

红外与激光工程. 2018, 47(9): 926002 <https://doi.org/10.3788/IRLA201847.0926002>

偏振参数最优重构的水下降质图像清晰化方法

Underwater degraded imagesharpener method based on optimal polarization parameter reconstruction

红外与激光工程. 2021, 50(6): 20200426 <https://doi.org/10.3788/IRLA20200426>

基于边界约束图像融合的光学字符识别算法研究

张 驯, 赵金雄, 白万荣, 赵 红

(国网甘肃省电力公司电力科学研究院, 甘肃 兰州 730000)

摘 要: 为了提高非合作目标区域中光学字符的识别能力, 增强电力铭牌、电网文本等信息采集的准确性, 设计了一种偏振图像与可见光图像同时采集并进行图像融合的光学字符识别系统。通过设置 0° 、 60° 和 120° 的偏振角对响应电压进行周期性调制, 从而获取有效信息区域的连通范围, 实现准确的边界约束。通过计算偏振角度标定参数, 设置边界条件的合理阈值, 为图像融合提供范围标准。实验分别测试了响应电压关于偏振角度和测试距离的函数曲线, 结果显示, 偏振角度周期性变化斜率为 $53.1 \text{ mV}/(^\circ)$ 。在 $0.5\sim 3.0 \text{ m}$ 范围内, 响应电压最大值为 241.7 mV , 最小电压为 18.5 mV , 并且三条响应曲线的单调性几乎一致。实验针对图像清晰度较差的电力铭牌目标进行测试, 结果显示, 模糊的原始图像在传统图像滤波与增强后, 对比度从 0.34 提升至 1.56 , 图像质量得到了一定的改善, 但仍有部分字符无法识别。而采用文中算法后, 对比度达到了 3.23 , 部分模糊字符也能有效识别。可见, 该系统适用于非合作目标的光学字符识别, 对低质量图像中光学字符识别具有很好的优化效果。

关键词: 偏振成像; 图像融合; 边界约束; 光学字符

中图分类号: TP391 **文献标志码:** A **DOI:** 10.3788/IRLA20220102

Research on optical character recognition algorithm based on boundary constrained image fusion

Zhang Xun, Zhao Jinxiong, Bai Wanrong, Zhao Hong

(Electric Power Research Institute of State Grid Gansu Electric Power Company, Lanzhou 730000, China)

Abstract: In order to improve the recognition ability of optical characters in non-cooperative target areas, and enhance the accuracy of information collection such as power nameplates and power grid texts, an optical character recognition system was designed that simultaneously collected polarized images and visible light images and performed image fusion. By setting the polarization angles of 0° , 60° and 120° , the response voltage was periodically modulated to obtain the connectivity range of the effective information region and achieve accurate boundary constraints. By calculating the polarization angle calibration parameters and setting reasonable thresholds of boundary conditions, a range standard was provided for image fusion. The function curves of the response voltage on the polarization angle and the test distance were tested in the experiment, and the results showed that the slope of the periodic change of the polarization angle was $53.1 \text{ mV}/(^\circ)$. In the range of $0.5\sim 3.0 \text{ m}$, the maximum value of the response voltage was 241.7 mV , the minimum voltage was 18.5 mV , and the monotonicity of the three response curves was almost the same. The experiment was carried out on the power

收稿日期: 2022-02-14; 修订日期: 2022-03-06

基金项目: 国网甘肃省电力公司科技项目 (52272220002U)

作者简介: 张驯, 男, 高级工程师, 硕士, 主要研究方向为信息安全、大数据与人工智能。

nameplate target with poor image definition. The results showed that after traditional image filtering and enhancement, the contrast ratio of the blurred original image was increased from 0.34 to 1.56, and the image quality was improved to a certain extent, but there were still some characters that could not be identified. After using this algorithm, the contrast ratio reaches 3.23, and some fuzzy characters can also be effectively recognized. It can be seen that the system is suitable for optical character recognition of non-cooperative targets, and has a good optimization effect on optical character recognition in low-quality images.

Key words: polarization imaging; image fusion; boundary constraints; optical characters

0 引言

在电网设备运检中,运检人员会以文本形式记录和以电子形式报告电力设备故障现象、消缺方法、运维检修措施等非结构化信息,这些信息不仅反映了设备运行状态的历史趋势,还可能蕴藏着丰富的潜伏性故障信息^[1]。传统设备文本数据主要依靠专家人员凭借经验对该设备的运行趋势进行判断^[2]。受人员经验、对设备的熟悉程度和文本描述方式等主观因素影响,易导致误检、漏检或误操作情况的发生,以至造成重大经济损失。随着变电站规模与数量的迅速增长,非结构化文本数据爆炸性的增长进一步凸显了电力文本数据处理、消化及再利用的问题,给日常电网设备运检带来了极大的安全隐患,不适应坚强智能电网的发展需求。由此光学字符识别技术得到了长足的发展。

现有的电力铭牌中包含信息复杂,结构类型多样,增加了识别难度^[3]。其中工作票及操作票以短文形式,主要包含运检人员周期性现场巡视、检修消缺、事后检修时记录的缺陷文字^[4]。上述不同类别的文本数据格式也不尽相同,其中包含各个时期的手写非结构文本、纸质电子文档等,并且由于环境差异巨大,图像大小、角度、色差等千差万别,故研究适用于复杂环境,具有高适应性的光学字符识别算法具有重要意义。

常见的光学字符图像识别系统主要包括:可见光图像识别技术^[5-6]、偏振图像识别技术^[7-8]和红外图像识别技术^[9-10]。王晓艺等人^[11]采用 OCR 技术对光学字符进行图像分割,局部待测字符灰度对比度从 0.48 提升至 1.39,增强了翻译文本字符的可识别性。Matas 等人^[12]采用区域特征提取的方式获取了不同阈值的二值化图像,实现了计算均值距离优于 0.09 个

连接像素距离。陈伟力等人^[13]采用红外光谱偏振技术完成涂层反射光谱的分析,相对光谱分辨率提升了 1 倍,可以获得更多的目标光谱信息细节。欧巧凤等人^[14]利用深度学习算法,对多种类型图像进行分类识别,识别精度优于 98%。周怡等人^[15]将模糊匹配算法应用于非标船牌的识别,提高了船牌识别概率。

文中采用图像融合技术增强复杂环境中光学字符的识别能力,提出了基于深度学习的识别系统架构,利用人工智能技术重新定义了光学字符识别“AI+OCR”,为实现复杂背景环境下光学字符识别提供一种新方法。

1 系统设计

系统整体设计结构如图 1 所示。

图中,处理系统控制调制模块使激光器发出周期性初始光信号,光波经偏振片产生固定相位的线偏振光。该光照射在待识别的电网文本表面产生反射,接

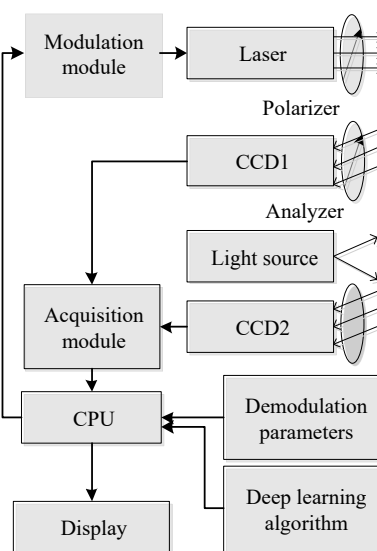


图 1 基于图像融合的光学字符识别系统

Fig.1 Optical character recognition system based on image fusion

收端经检偏器完成对具有一致偏振态回波光的检测,并成像于 CCD1 上。与此同时,可见光光源同时照射待测区域,由 CCD2 获取对应位置的可见光二维图像。通过采集模块统一坐标系后,由 CPU 完成图像融合。系统通过深度学习算法和解调参数实现对光学字符的再现。

2 图像融合预处理

由于数据来源有两个,一个是偏振强度图像,另一个是可见光图像,所以在文中的系统中,将偏振强度图像的连通包络区域作为光学字符识别的限定条件,然后在此基础上完成不同权值的配准,从而提高可见光图像中光学字符的可辨识度。

2.1 基于偏振图像的边界条件

采用 Stocks 法^[11]解算偏振态,当入射偏振角可以预设时,则其矢量的强度值可以由初始光源的功率分布曲线获得。该 Stocks 矢量可写成:

$$S = [I \quad Q \quad U \quad V]^T \quad (1)$$

式中: I 表示光强; Q 和 U 分别表示偏振分量值; V 表示圆偏振分量。则在 CCD1 上,其穆勒矩阵有:

$$S_{\text{CCD1}} = M \cdot S_i \quad (2)$$

式中: S_{CCD1} 为 CCD1 上采集到的斯托克斯矢量; S_i 为初始照射到目标上的斯托克斯矢量; M 为穆勒矩阵。

由于光学字符与没有光学字符的位置表面存在差异,一种是由于书写导致的表面反射率的不同,一种是电力贴签上字符位置本身是由不同材料印刷上的,故根据不同偏振强度值可以对测试区域的图像进行范围划分,并赋予不同的权值,从而提高有效信息的提取能力。

由于待测电力铭牌等均具有一定的一致性,所以偏振光偏振态不需要大范围扫描,仅需采用几个固定偏转角的偏振光即可,文中系统中采用 0° , 60° 和 120° 。不同偏振方向的响应结果可以通过均匀化校正得到,若设 0° 时响应为基准,则另外两个偏振角度的标定参数有:

$$k_{60^\circ} = \frac{V'_{0^\circ}}{V'_{60^\circ}}, \quad k_{120^\circ} = \frac{V'_{0^\circ}}{V'_{120^\circ}} \quad (3)$$

式中: V'_{60° 表示偏振角为 60° 时 CCD1 的响应电压均值; V'_{120° 表示偏振角为 120° 时 CCD1 的响应电压均

值。通过与 0° 时进行比较从而可以确定两个偏振角度的标定参数。

2.2 图像融合

偏振图像在文中系统中的主要作用是提供准确的图像边界信息,所以采用小波不变矩提取偏振图像中的特征范围可以大幅减少全图像数据分析的压力,从而保证测试系统的识别速度。设图像为 $g(x,y)$, 则其第 $s+t$ 阶的原点矩可表示为:

$$M_{st} = \iint x^s y^t g(x,y) dx dy \quad (4)$$

式中: x 、 y 分别表示二维图像坐标值; s 、 t 分别表示 x 方向与 y 方向上的阶数。因为测试过程中,铭牌类别各异,尺寸相差较大,若如果采用直角坐标进行图像对齐会使增加多余的计算量,故将其转化为极坐标系,可以提升图像融合效率。由直角坐标系与极坐标系的关系:

$$x = r \cos \phi, \quad y = r \sin \phi \quad (5)$$

公式 (4) 可以表示为:

$$M_{st} = \iint r^{s+t} (\cos \phi)^s (\sin \phi)^t g(r, \phi) r dr d\phi \quad (6)$$

由此得到偏振图像中表达测试文字区域的原点矩参量,可以实现对特征范围的限定,再将该限定范围内的图像进行增强处理,就能快速高效地获取电力铭牌等上的有效信息了。

3 光学字符识别算法

3.1 特点分析

由于电力铭牌等上的字符往往由文本、数字甚至是特殊符号构成,所以不能简单地采用文字识别算法。电力系统中专业术语多,缩略表达也十分常见,同时,有效信息的长度、记录人员主观表达方式等都对含义有一定影响,所以获取数据必须尽可能地保证全面,减小因信息缺失造成的信息错误。综上所述,光学字符识别算法的构建需要综合考虑多种情况,故在模型设计中需要给出适用于多种状态分类的计算方法。

3.2 模型构建

传统光学符号识别常采用诸如基于 Bi-LSTM 的特征识别算法、基于胶囊网络的聚类分析算法和基于滑动窗的预测识别算法。在文中系统中,由于需要分

析的变量不多,但变量类型较复杂,所以相比之下,采用 Bi-LSTM 算法比较合适。在 Bi-LSTM 算法中,主要包括输入模块、遗忘模块、输出模块及存储模块。设当前输入词的向量值为 x_t , 而其隐状态的两个时刻分别是 h_{t-1} 和 h_t 时,则 Bi-LSTM 算法中的特征参数可表示为:

$$\begin{cases} h_t = \sigma(A) \tanh\{\sigma(A)[\tanh(A) + c_{t-1}]\} \\ A = W_i[x_t] + U_i h_{t-1} + V_i c_{t-1} + b_i \end{cases} \quad (7)$$

式中: σ 表示 Sigmoid 函数; h_t 表示函数输出,其前向输出与后向输出的对位相加可以计算得到; t 表示当前时刻; $t-1$ 表示上一个时刻; i 表示函数计算的第 i 个数据; A 为计算过程中的中间变量; $W_i[x_t]$ 为第 i 个数据在 t 时刻向量值 x_t 对应的测试值; U_i 和 V_i 分别为相对 W_i 测试值的补偿参数; c_{t-1} 为 $t-1$ 时刻的偏振补偿参数; b_i 为修正常数。

上式中将测试值关于测试值强度、时刻变化以及偏振补偿的影响统一到了一个模型中,从而能够在光学符号识别的过程中同时选择符合偏振边界及时序变化规律的图像区间。

3.3 算法实现

在完成基于偏振特征的边界条件约束和光学字符识别模型的基础上,按照参数计算顺序与数据分析流程就能获取准确的电力铭牌等测试信息。算法步骤如下:

(1) 根据测试数据类型分别进行处理,结合测试时偏振镜的偏振角度,对偏振数据进行两个偏振角度标定参数的计算,即计算得到 k_{60° 和 k_{120° 。通过比较不同偏振角度的比值,完成对边界范围的选取;

(2) 对获得的二维可见光图像进行原点距计算,从而为图像目标识别提供依据,在此过程中,将原有不同大小的直角坐标系统一成原坐标系进行表达,从而避免基于图像原点的测试字符范围偏差过大的问题,求解 M_s , 获得图像融合上下限及对齐角度;

(3) 将偏振边界范围作为限定条件代入图像解算模型,从而大幅降低图像融合数据处理的运算量,计算得到测试值 $W_i[x_t]$ 、 U_i 和 V_i ;

(4) 根据测试值之间的关系计算光学字符识别的中间变量 A , 将计算得到的 A 值代入公式 (7) 完成时刻特征参数的计算;

(5) 通过特征参数 h_t , 循环迭代筛选图像区域,完成所有图像中目标字符集字符串的识别。

整体识别程序流程如图 2 所示。

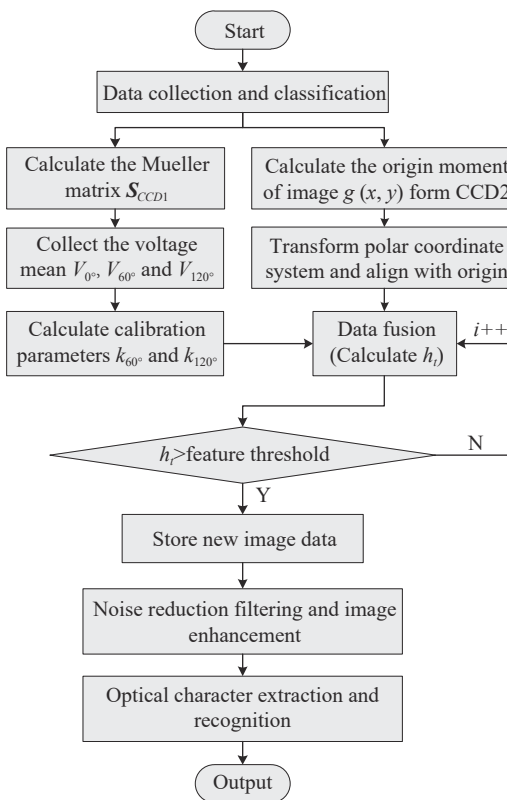


图 2 光学字符识别算法程序流程图

Fig.2 Flow chart of optical character recognition algorithm

4 实验

4.1 偏振角度标定实验

根据已搭建的偏振识别系统,对选择的三个偏振角度 0° 、 60° 和 120° 进行测试,从而为边界选取提供量化对比依据。实验通过 555.0 nm 激光分别照射探测器,获取不同偏振角度条件下的响应电压,并且与无检偏镜对比,测试曲线如图 3 所示。

图 3 中,当没有检偏器时,回波强度信号完全没有变化,即可以理解成采用可见光照明的效果。当采用不同角度偏振后,其响应电压强度变化关系基本呈周期性变化,周期角度为 180° ,三条测试曲线的相位差为 60° 。最大响应电压分别是 4.77 V 、 4.75 V 和 4.78 V ,与无检偏镜时测试得到的 4.79 V 相近。最小响应电压分别是 0.08 V 、 0.11 V 和 0.09 V , 180° 之前的相位由于是初始相位,所以角度对准得更好,故其

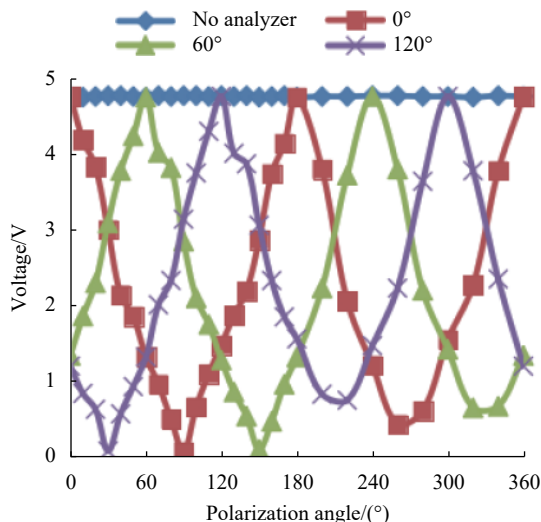


图 3 偏振角度与电压之间的函数关系

Fig.3 Function relationship between polarization angle and voltage

最小值普遍优于后面的测试值。响应电压变化斜率为 53.1 mV/(°), 可见在偏振角度位置合适时可使响应电压达到最大值。由于采用了偏振调制, 可以根据待测光学字符的表面反射程度选择合适的偏振角度, 从而实现对待定区域的高反射效果, 增强光学字符的识别概率。

4.2 距离标定实验

测试过程中, 除了偏振角度会对响应电压造成影响, 测试距离同样对测试强度影响明显, 故需要对不同测试距离条件下的电压响应值进行测试标定, 才能在后期图像融合中选择合适的图像增强补偿系数。由于测试系统与电力铭牌、标志符等位置不会太远, 故测试距离范围设定在 0.5~3.0 m 之间, 每 0.1 m 采集一次响应电压变化关系, 并同时记录无偏振镜、0°、60°和 120°偏振镜的响应值, 测试结果如图 4 所示。

图中, 随着测试距离的增大, 无论是否有偏振镜, 其响应电压都会随之降低。不同偏振角对应的响应电压变化趋势几乎一致, 0.5 m 处最大响应电压分别为 234.8 mV、241.7 mV 和 238.0 mV, 3.0 m 处最小响应电压分别为 18.8 mV、18.5 mV 和 19.2 mV。虽然非偏振条件下的测试响应更大, 但是偏振可以获取调制信息, 而从振幅强度和单调性上看, 三个偏振角度对应的测试曲线与非偏振的测试值相近。由测试曲线的单调性可知, 不同测试距离对应的响应电压是唯一的, 所以可以通过该曲线的标定数据完成对不同偏振角照射条件下的距离的区分, 故可实现相同距离上光

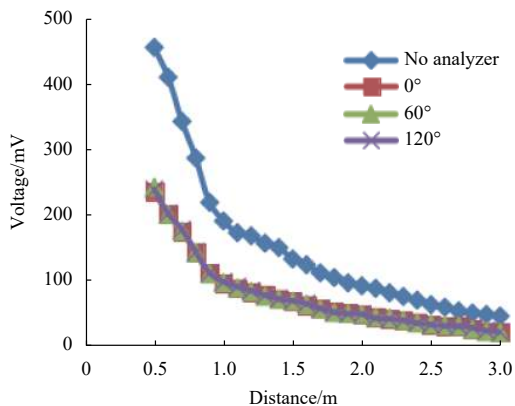


图 4 测试距离与电压之间的函数关系

Fig.4 Function relationship between test distance and voltage

学字符区域的选取, 验证了文中算法可区分不同距离响应电压的可行性。

4.3 电力铭牌字符识别实验

为了验证文中系统可以提高光学字符的识别能力, 针对非合作电力铭牌进行二维图像采样, 分别采用直接图像获取、图像滤波增强以及基于偏振边界条件下的图像增强三种识别方式。为了更准确地表达光学字符的识别效果, 仅提取测试图像中包含有效信息的区域进行对比, 测试距离 1.0 m, 非合作照明补光 (采用照相机正常闪光灯照相模式), 测试与优化后图像如图 5 所示。

如图 5(a) 所示, 原始图像由于受到测试位置与环境的影响, 虽然有照明补光, 但是其图像质量仍很差, 其中铭牌上的主要参数有些无法识别; 如图 5(b) 所示, 对原始图像进行去噪滤波再进行图像增强就能得到较清晰的光学字符影像, 其中型号、额定功率、编号等信息基本可以识别, 个别光学字符不清晰, 存在误识别问题; 如图 5(c) 所示, 在采用偏振图像进行边界分类后, 电力铭牌上具有明显边界范围的地方都被分类为强保留区域, 故从标识字符到信息字符的边界都得到了大幅增强, 其中编号中的最后一个字符“8”和额定功率中的第一个字符“5”都有了明显的增强。提取有效字符位置区域灰度值均值与其相邻区域像素灰度值均值进行对比度计算, 图 5(a)、(b) 和 (c) 的结果分别是 0.34、1.56 和 3.23, 采用算法优化后待测字符像素对比度提升了近一个数量级。由此可见, 采用本系统对于非合作目标光学字符识别具有一定的优势。



(a) 原始图像
(a) Original image



(b) 传统图像增强
(b) Traditional image enhancement



(c) 基于偏振边界的图像融合
(c) Image fusion based on polarization boundary

图 5 光学字符识别测试结果对比

Fig.5 Comparison of optical character recognition test results

5 结 论

文中搭建了偏振图像与可见光图像同时获取的光学字符识别系统,并基于此设计了基于偏振图像边界约束的图像融合算法。算法通过不同偏振角度完成响应电压调制,再利用偏振角度标定参数获得融合图像范围。实验测试了偏振角度、测试距离对响应电压的影响,结果显示:偏振角度调制曲线均匀,测试距离的电压响应曲线单调且具有很高的重叠性,验证了系统算法的可行性。最后,通过与原始图像、传统图像增加算法的对比,验证了文中算法的识别效果最佳。总之,文中系统可以更好地用于非合作光学字符目标的识别。

参考文献:

[1] Deng C, Pan L, Wang C, et al. Performance analysis of ghost imaging lidar in background light environment [J]. *Photonics Research*, 2017, 5(5): 431-435.
 [2] Yu H, Li E, Gong W, et al. Structured image reconstruction for

three-dimensional ghost imaging lidar [J]. *Optics Express*, 2015, 23(11): 14541-14551.
 [3] Wang X, Shao Y M, Yang B, et al. Target tracking method based on infrared and laser lidar image fusion [J]. *Infrared Technology*, 2019, 41(10): 947-955. (in Chinese)
 [4] Ling Y, Gu G, He W, et al. Adaptive target profile acquiring method for photon counting 3-D imaging Lidar [J]. *IEEE Photonics Journal*, 2016, 8(6): 1-10.
 [5] Tian Z, Cui Z, Zhang L, et al. Control and image processing for streak tube imaging lidar based on VB and MATLAB [J]. *Chinese Optics Letters*, 2014, 12(6): 67-70.
 [6] Zhong H, Yuan Y, Wang J F, et al. Anchor box optimization for object detection [EB/OL]. (2018-12-02) [2022-02-14]. <https://arxiv.org/pdf/1812.00469.pdf>.
 [7] Xie L L, Liu Y L, Jin L W, et al. DeRPN taking a further step toward more general object detection[C]//Proc of the 33rd Association for the Advancement of Artificial Intelligence. Menlo Park, CA: AAAI Press, 2019: 9046-9053.
 [8] Yang T, Zhang X Y, Li Z M, et al. MetaAnchor: Learning to detect objects with customized anchors[C]//Proc of the 31st Annual Conference on Neural Information Processing Systems. New York: Curran Associates Press, 2018: 318-328.
 [9] Wang J Q, Chen K, Yang S H, et al. Region proposal by guided an Choring [C]//Proc of IEEE Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2019: 2965-2974.
 [10] Ma K, Liu S H, Bai X, et al. DocUNet: Document image unwarpping via a stacked U-net[C]//Proc of IEEE Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2018: 4700-4709.
 [11] Wang X Y, Gao T T. Translation optimization based on OCR optical character recognition [J]. *Laser Journal*, 2020, 41(12): 156-160. (in Chinese)
 [12] Matas J, Chum O, Urban M, et al. Robust wide-baseline stereo from Maximum stable extreme regions [J]. *Image and Vision Computing*, 2004, 22(10): 761-767.
 [13] Chen W L, Xu W B, Wang S H, et al. Research on coating materials detection and recognition based on infrared spectral polarization degree contrast [J]. *Infrared and Laser Engineering*, 2020, 49(6): 20190445. (in Chinese)
 [14] Ou Q F, Xiao J B, Xie Q Q, et al. Multi-target detection and recognition for vehicle inspection images based on deep learning [J]. *Journal of Applied Sciences*, 2021, 39(6): 939-951. (in Chinese)
 [15] Zhou Y, Zhu Q R, Xie H C, et al. Non-standard ship identification characters detection based on target detection and fuzzy matching [J]. *Laser & Infrared*, 2021, 51(11): 1526-1530. (in Chinese)