

doi: 10.11823/j.issn.1674-5795.2022.03.03

基于深度学习的合作目标靶球检测

王国名^{1,2}, 郝灿^{1,2}, 石俊凯¹, 高超^{1,2}, 王博¹, 周维虎^{1,2*}, 高豆豆^{1*}

(1.中国科学院微电子研究所,北京 100094;2.中国科学院大学,北京 101408)

摘要: 为了解决复杂场景下激光跟踪仪对合作目标靶球的精确识别难题,提出了基于深度学习的合作目标靶球高效检测方法。首先分析了合作目标靶球的图像特征,然后采用改进的YOLOv2模型,针对合作目标靶球多尺度与小目标占比多的特点,提出了一种基于注意力机制的改进方法,同时为提高网络模型对复杂背景的抗干扰能力,提出了一种数据增强方法。测试结果表明,所提出的基于注意力机制与数据增强的改进YOLOv2模型对复杂背景的抗干扰能力较强,且对合作目标靶球的检测精度有显著提高,在合作目标靶球测试集上的检测准确率达到92.25%,能够有效满足激光跟踪仪在大型装置精密装配过程中的目标检测精度需求。

关键词: 激光跟踪仪; 合作目标靶球检测; 深度学习; YOLOv2

中图分类号: TB96

文献标识码: A

文章编号: 1674-5795(2022)03-0016-07

Object detection of cooperative target based on deep learning

WANG Guoming^{1,2}, HAO Can^{1,2}, SHI Junkai¹, GAO Chao^{1,2}, WANG Bo¹,

ZHOU Weihu^{1,2*}, GAO Doudou^{1*}

(1.Institute of Microelectronics of the Chinese Academy of Sciences, Beijing 100094, China;

2.University of Chinese Academy of Sciences, Beijing 101408, China)

Abstract: In order to improve the detection accuracy of cooperative target ball used for the precision assembly of large-scale devices by laser tracker in complex scenes, an efficient cooperative target ball detection method based on deep learning is researched. Firstly, the image features of the cooperative target are analyzed. Then, by using the improved YOLOv2 model, an improved method based on attention mechanism is proposed aiming to the cooperative target characteristics of multi-scale and large proportion of small targets. In order to improve the anti-interference ability of the network model in complex background, a method of data enhancement is also proposed. The test result shows that the proposed improved YOLOv2 network based on attention mechanism and data enhancement has strong anti-interference ability against complex background and significantly improves the detection accuracy of cooperative target ball. The detection accuracy on the cooperative target test set has reached 92.25%, which meets the target detection accuracy requirements of laser tracker in the large equipment precision assembly.

Key words: laser tracker; cooperative object detection; deep learning; YOLOv2

0 引言

随着我国先进制造和智能制造业的快速发展,

对所需高端装备的生产工装测量、精密装配检测及整机精度检测的精度要求越来越高,对以激光跟踪仪为代表的大尺度空间坐标精密测量仪器的

收稿日期: 2021-08-31; 修回日期: 2022-05-10

基金项目: 国家重点研发计划(2019YFB2006100)

引用格式: 王国名, 郝灿, 石俊凯, 等. 基于深度学习的合作目标靶球检测 [J]. 计测技术, 2022, 42 (3): 16-22.

Citation: WANG G M, HAO C, SHI J K, et al. Object detection of cooperative target based on deep learning [J]. Metrology and measurement technology, 2022, 42 (3): 16-22.



需求也越来越迫切^[1-3]。在飞机、船舶等大型装置的精密装配过程中,激光跟踪仪如何在复杂生产环境下对合作目标靶球进行快速精确识别捕获,从而精确引导激光自动对准合作目标,是提升大型装置精密装配效率、提高断光续接精度亟需解决的难题^[4]。为了解决目标的自动识别与自动测量问题,张滋黎等人^[5]提出了一种基于视觉引导的激光测量系统,实现了激光的自动对准;中国科学院大学^[6]提出了一种大目标高精度位置姿态测量技术,利用单目相机获取合作目标的中心位置,准确引导激光测距仪自动瞄准合作目标的中心位置,通过参数运算得到待测目标的空间位姿,从而实现高精度自动引导测量。

近年来,基于深度学习的机器视觉目标检测方法在精密检测领域得到广泛应用,将深度学习用于合作目标检测,可极大地提高激光跟踪仪对其合作目标靶球的瞄准精度。具有代表性的深度学习方法有基于区域推荐的目标检测方法(如Faster RCNN^[7])和基于回归的目标检测方法(如YOLO^[5])。基于区域推荐的目标检测方法分两个步骤,首先预测可能存在目标的区域,然后进行目标的分类与边界定位,因此检测速度较慢,难以满足实时检测的需求。

2016年,Redmon等人^[8]首创YOLO目标检测方法,使用单个卷积神经网络直接对目标进行位置回归与类别预测,检测速度可达到同时期两步骤算法的5倍,突破了目标实时检测的难题。但是,该方法在显著提高检测速度的同时模型检测精度有所下降,尤其对小目标的检测精度较低,定位目标准确性差。针对此问题,Liu等人^[9]提出了一种SSD(Single Shot MultiBox Detector)方法,在不同尺度特征图上分别生成目标的先验框,提高了模型对不同尺度目标的检测精度。Redmon和Farhadi^[10]提出的YOLOv2方法借鉴了SSD对多尺度特征的处理策略,对小尺寸目标的检测精度有所提高,且其检测精度能够满足大部分应用场景。然而,该方法仍然采用特征提取能力较弱的Darknet19基础网络,目标检测精度有待进一步提高。

深度学习方法利用卷积神经网络丰富提取图像特征检测目标,能够满足激光跟踪仪不同工作距离条件下多尺度目标检测的需求。其中,YO-

LOv2方法使用单个卷积神经网络直接对目标进行分类与回归预测,检测速度快,满足激光跟踪仪实时检测合作目标靶球的需求。但是YOLOv2的基础网络特征提取能力较弱,针对小目标的定位能力较差,且复杂环境条件下的抗干扰能力较差,制约了合作目标靶球的检测精度。通过改进基础特征提取网络,可进一步提高YOLOv2网络对多尺度目标的提取能力,从而提高检测精度。

针对激光跟踪仪跟踪的合作目标靶球形状单一、纹理细节较少且场景复杂的问题,改进基础特征网络,提出一种基于YOLOv2单阶段目标检测模型的靶球检测方法,能够有效克服复杂背景的干扰,提高合作目标靶球的检测精度,从而提升激光跟踪仪对合作目标的瞄准精度。

1 基于YOLOv2卷积神经网络的合作目标靶球检测方法

1.1 YOLOv2网络结构原理

Joseph Redmon在2016年提出了YOLOv2网络结构^[10],图像输入网络后被划分为 $S \times S$ 个网格,每个网格通过卷积神经网络后预测生成 K 个边界框。边界框包含目标的置信度为 $P_r(\text{Object})$,预测边界框与真实边界的交并比为 $IoU_{\text{pred}}^{\text{truth}}$,目标属于第 i 类的条件概率为 $P_r(\text{Class}_i/\text{Object})$,则每个预测框的类别置信度表示为

$$\begin{aligned} P_r(\text{Class}_i/\text{Object}) * P_r(\text{Object}) * IoU_{\text{pred}}^{\text{truth}} \\ = P_r(\text{Class}_i) * IoU_{\text{pred}}^{\text{truth}} \end{aligned} \quad (1)$$

式(1)表征了目标的预测边框与真实边框重合程度以及目标属于第 i 类的概率。

大型装置精密装配时,合作目标的运动范围广且随机性强,背景环境复杂,将YOLOv2直接用于合作靶球实时检测仍存在一定挑战。YOLOv2候选框所能匹配到的正负样本比例极为悬殊,而YOLOv2的“Fine-Grained Features”方法偏重于融合不同层的全局特征,融合时会损失部分多尺度特征图的局部细节信息,造成对小尺寸目标的检测精度下降,制约了合作目标检测精度。

1.2 针对靶球尺度变化及小目标检测的改进网络

除了模型与各项参数外,训练数据集的样本质量也影响着深度学习网络的最终检测性能^[11-12]。若待测样本极易被模型检测,则其被认

为是简单样本；若待测样本较难被模型检测，则其被认为是困难样本。相关研究表明：在模型训练困难样本时，会产生较大的梯度变化，通过降低简单样本的权重，增大困难样本在训练中的权重，使其对损失的贡献变大，能够有效提高模型的优化效率。

文章提出了一种结合注意力机制框架结构与困难样本挖掘的方法，通过注意力机制来自适应提取合作目标靶球周围的有用信息。注意力机制模块从多个感受野提取到大量上下文信息，并自适应提取显著对象周围的有用信息。同时使用了更深度化的特征提取网络，提取更丰富的全局语义信息与局部定位信息，从而得到更好的检测与

定位精度。基于注意力机制与困难样本挖掘的改进YOLOv2模型网络图如图1所示，使用RestNet50替换了原YOLOv2中的Darknet19特征提取网络，并增加了注意力机制（AM）模块以提升模型对困难样本的挖掘能力。在图像输入特征提取网络前对其进行切片（Slice）操作，使用最近邻下采样（Focus）将原本只有RGB三通道的输入图像扩充到12个通道，经过卷积组件（CBL）和残差组件（Res Unit）处理后，进行特征融合（Concat），得到包含丰富特征且无信息损失的特征图，在输出层利用注意力机制获取有用的目标信息。相较标准的YOLOv2网络，新的特征图具有更强的小目标检测能力，且在复杂背景下的鲁棒性更强。

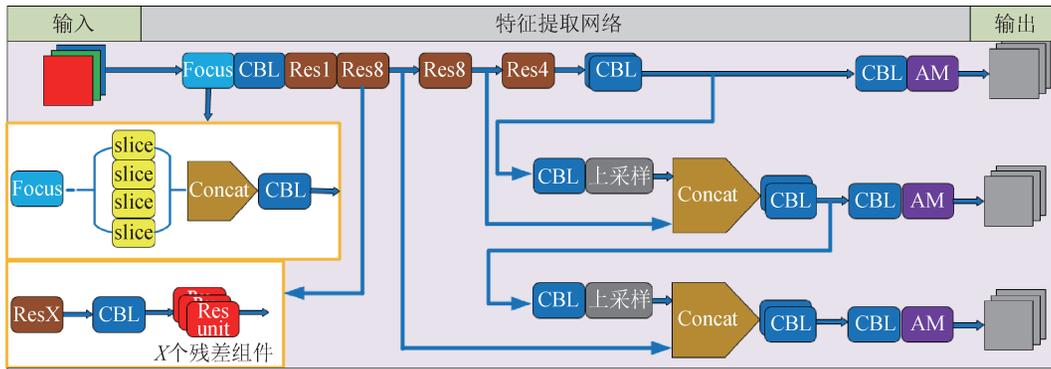


图1 基于注意力机制与困难样本挖掘的改进YOLOv2模型网络图

Fig.1 Improved YOLOv2 network based on attention mechanism and difficult sample mining

2 合作目标靶球数据集分析与数据增强

2.1 合作目标靶球数据集的建立

构建的合作目标靶球数据集使用面阵CCD相机采集，使用PASCALVOC标准数据集格式标注，部分样本图像如图2所示，其中(a)，(b)，(d)为不同场景的图像，(b)，(c)为同一场景不同角度的图像，标注目标为合作目标靶球。合作目标靶球在图像中接近圆形，在构建数据集时用圆形瓶盖、圆形器件等构建靶球的复杂背景。

数据集包含1056张图像。每张图像大小为640×480 pixels，每幅图像仅包含一个合作目标靶球，其最小边界框的左上角与右下角坐标作为真实值保存为标签。在真实应用场景中，合作目标靶球可能出现在图像的任意位置，数据集中合作目标靶球的位置分布如图3(a)所示，数据集中

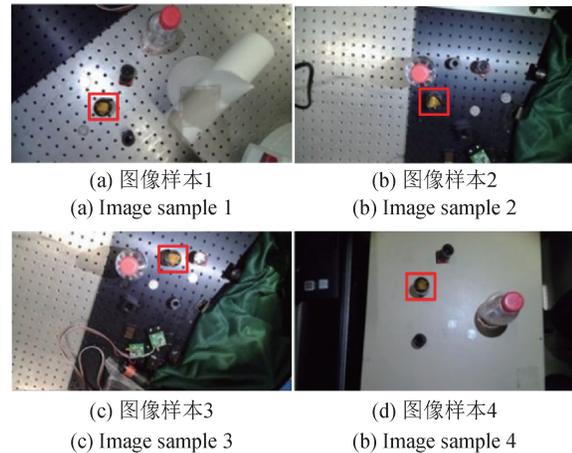


图2 部分合作目标靶球数据集图像

Fig.2 Sample images of the cooperative target ball dataset

合作目标靶球在图像中的面积占比分布如图3(b)所示，可以看出长宽比小的区域样本密集，小目标在整个数据集中占比较高。

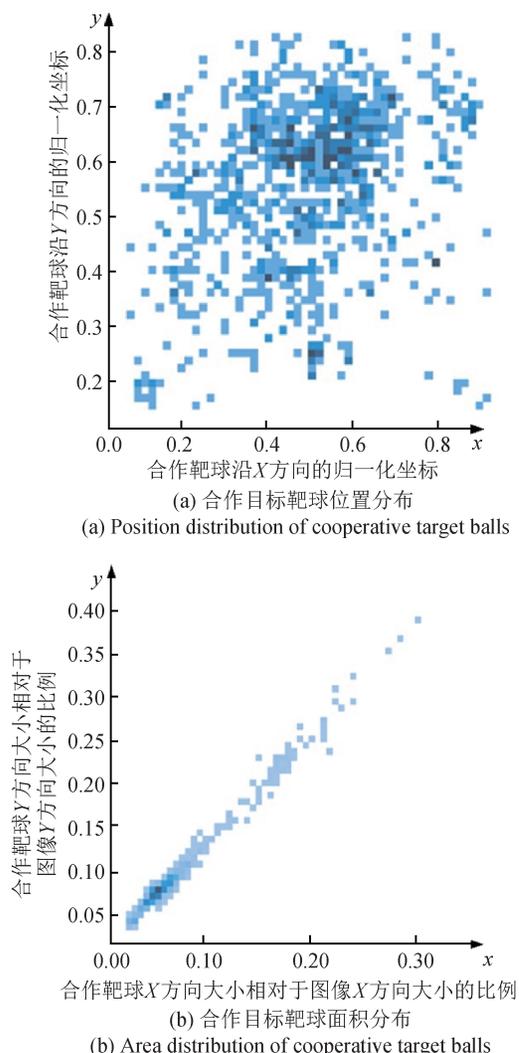


图3 合作目标靶球分布图

Fig.3 Distribution of cooperative target balls

2.2 基于复杂背景的数据增强

深度学习目标检测模型的检测性能一定程度上受数据集质量的影响。在实际应用中背景极为复杂，且合作目标靶球在图像中的大小会发生变

化，为了增加模型对强背景干扰的鲁棒性并提高图像中任意大小合作目标靶球的检测精度，需要针对小目标进行数据增强^[13]。

实验中发现，使用 Mixup^[14] 数据增强方法不能解决复杂背景造成的检测精度下降的问题。由于合作目标靶球的结构外形单一，形状近似的背景目标极易对检测造成干扰，且训练样本中小目标数据较少，采用测试集检测小目标样本时效果较差。因此，本文提出了一种图像拼接的数据增强方法：从原始数据集随机选取4张图像，对4张图像分别进行翻转、旋转、缩放操作后，拼接成一幅图像，对于缩小图像进行0像素填充，扩充图像至原图像大小，对于放大图像随机截取包含合作目标靶球的和原图大小相同的区域，如此就形成了新的增强数据集。拼接增强后的图像增加了合作目标背景复杂度，从而提高模型对复杂背景干扰的鲁棒性。且由于4张图像拼接减小了训练样本合作目标靶球的尺寸，从而有效地提高了模型对小尺度合作目标靶球的检测精度。采用数据集中随机4幅图像进行图像拼接后生成的数据增强图像如图4所示，其中O为目标靶球。该方法增强了原数据集中图像的背景复杂度，有效地增加了检测模型对复杂环境的鲁棒性，以拼接图像作为训练样本，使合作目标靶球的检测准确率得到显著提高。

3 实验与结果分析

3.1 合作目标靶球检测的评价指标

合作目标靶球检测使用平均准确率 AP (Average Precision) 和每秒检测帧数 FPS (Frames Per

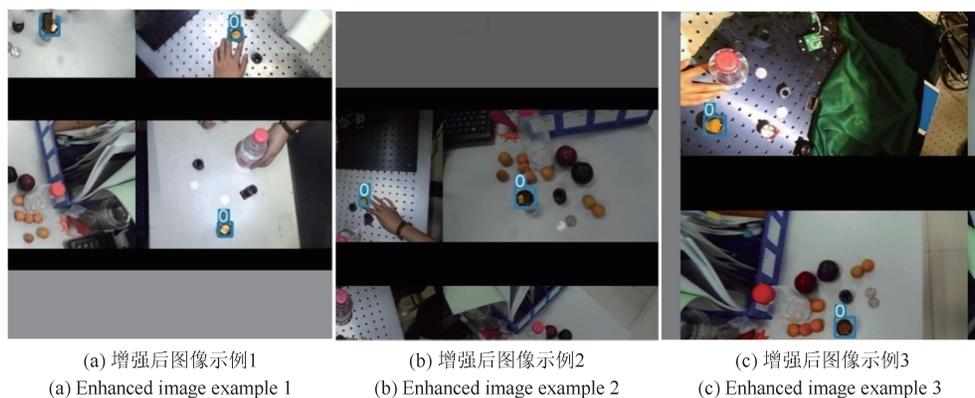


图4 部分经过数据增强后的合作目标靶球数据集图像

Fig.4 Sample images of the cooperative target ball dataset by stitching enhancement

Second) 作为评价指标。如图5所示, 内部矩形为预测为合作目标靶球的样本数量, 外部矩形为所有被检测为非合作目标靶球的样本数量。TP (True Positives) 为预测与实际均为合作目标靶球的样本数量, FP (False Positives) 为预测为合作目标靶球但实际为干扰目标的样本数量, FN (False Negatives) 为未检出合作目标靶球的样本数量, TN (True Negatives) 为预测为非合作目标且实际为非合作目标靶球样本数量。

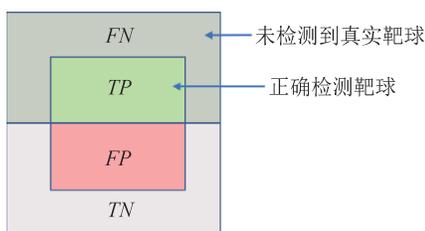


图5 预测合作目标靶球与真实合作目标靶球的TP, FP, FN和TN的定义

Fig.5 Definitions of TP, FP, FN and TN for predicted cooperative target balls and real cooperative target balls

查准率 p (Precision) 和查全率 r (Recall) 可分别表示为

$$p = \frac{TP}{TP + FP} \quad (2)$$

$$r = \frac{TP}{TP + FN} \quad (3)$$

AP为合作目标靶球PR曲线 (Precision-Recall Curve) 所包含的面积。使用11点法^[15] 计算各个类别的AP, 计算公式如式(4)所示。

$$AP = \frac{1}{11} \sum_{r \in \{0, 0.1, \dots, 1\}} p(\tilde{r}) \quad (4)$$

式中: $p(\tilde{r})$ 为计算得到 $r = [0, 0.1, 0.2, 0.3, \dots, 1]$ 时曲线下方的面积。

3.2 合作目标靶球检测实验结果

按照 8 : 1 : 2 的比例将原始数据集分为训练集、验证集与测试集。采用2.2节所述的方法对原始数据集进行数据增强。采用未参加训练的207张原数据集中的图像进行算法性能测试, 其中小目标测试样本约占总测试样本数量的40%。实验所使用的高性能图像工作站配置为 Intel Core I7 v4 3.3 GHz CPU, 128.00 GB 内存, RTX2080 Ti GPU, 11.00 GB GPU 显存。合作目标靶球检测实验在

Windows10 系统 Pycharm 平台和 Visual Studio 2019 平台上进行, 所有程序均采用python语言和C/C++语言编程实现; 程序所使用的深度学习框架为PyTorch (C语言对应为LibTorch)。模型训练过程中的动量设置为0.9, 训练批次大小为4, 权值衰减系数为 1×10^{-4} , 训练初始学习率为 5×10^{-3} , 最大迭代次数为30000, 在迭代进行到5000次和12000次时, 学习率分别减小至 5×10^{-4} 和 5×10^{-5} 。

为比较本文提出的结合注意力机制和困难数据挖掘的改进YOLOv2的检测性能, 分别对FasterRCNN、标准YOLOv2和本文提出的改进YOLOv2三种模型采用原始数据集和数据增强后数据集进行模型训练, 并采用同一测试集进行合作目标靶球检测, 6种测试结果见表1。由实验结果可知, 相较原始数据集, 经数据增强后的数据集检测精度有了显著提高。本文提出的改进YOLOv2方法的检测准确率为92.25%, 相较FasterRCNN模型检测准确率提高了5.79%, 且检测速度接近FasterRCNN模型的5倍, 能够有效满足合作目标靶球高精度检测要求。

表1 不同模型合作目标靶球检测结果

Tab.1 Detection results of cooperative target balls by different models

目标检测模型	平均准确率 AP/%	检测速度FPS/ (frame·s ⁻¹)
FasterRCNN	76.53	2.4
FasterRCNN+数据增强	86.46	2.4
YOLOv2	67.84	10.6
YOLOv2+数据增强	79.26	10.6
改进YOLOv2	80.43	9.7
改进YOLOv2+数据增强	92.25	9.7

利用标准的YOLOv2与改进的YOLOv2对同一数据集中不同图像中合作目标靶球进行对比检测试验, 结果如图6和图7所示, 同时在3个场景条件下进行检测, 标准YOLOv2模型在复杂背景下对小目标存在误检和漏检, 而本文提出的改进的YOLOv2方法在复杂背景下的抗干扰性能较好, 能够保持较高的小目标检测准确率, 相较标准YOLOv2方法具有明显优势。



图6 标准YOLOv2模型的不同图像中目标靶球的检测结果

Fig.6 Detection results of target ball in different images of classical YOLOv2 network

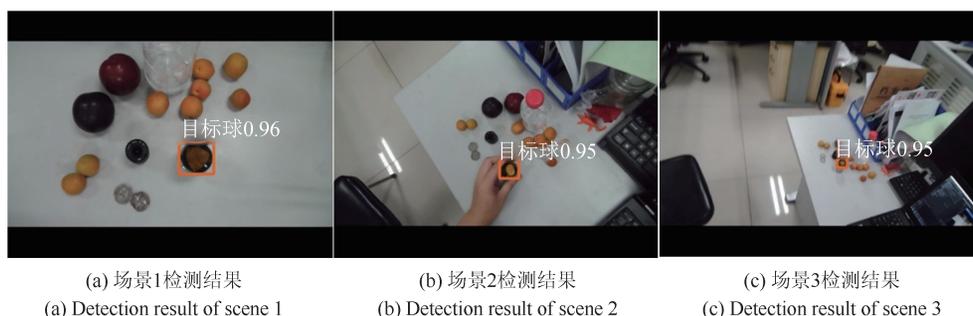


图7 改进YOLOv2模型的不同图像中目标靶球的检测结果

Fig.7 Detection results of target ball in different images of improved YOLOv2 network

4 结论

为解决激光跟踪仪在大型装置精密装配过程中对合作目标瞄准精度较低的问题,探寻基于深度学习的目标检测方法,以提升复杂背景下合作目标靶球的检测准确率与检测速度。分析了实际应用中检测距离变化而使靶标在图像中尺寸不同,导致YOLOv2深度卷积神经网络算法模型对合作目标靶球检测精度较低的现象,提出引入基于注意力机制与困难样本挖掘的方法提升多尺度目标检测精度;针对复杂背景条件下,合作目标靶球外形单一,纹理细节较少的特点,提出了基于拼接图像对数据集进行数据增强的方法,增加复杂背景和小目标样本的数量,有效提高了模型的检测准确率。实验结果表明,本文提出的改进YOLOv2模型和数据增强方法能够显著提高复杂背景下小目标合作目标靶球的检测精度,适用于远距离的合作目标跟踪,能够满足激光跟踪仪的实际应用要求。

参考文献

[1] 孙安斌, 曹铁泽, 王继虎, 等. 高端装备大型零部

件几何尺寸测量技术现状及趋势[J]. 计测技术, 2021, 41(2): 41-50.

SUN A B, CAO T Z, WANG J H, et al. Technological development trends of geometric dimension measurements of large parts in the high-end equipment[J]. Metrology and measurement technology, 2021, 41(2): 41-50. (in Chinese)

[2] 金挺, 梁琦, 王瑛辉. 激光标线仪校准方法的探讨[J]. 计测技术, 2020, 40(4): 23-26.

JIN T, LIANG Q, WANG Y H. Discussion on the calibration method of laser marker [J]. Metrology and measurement technology, 2020, 40(4): 23-26. (in Chinese)

[3] 林恬, 张学涛, 马骊群. 基于数字化设计制造协同飞机计量保障体系架构探讨[J]. 计测技术, 2018, 38(2): 59-63.

LIN T, ZHANG X T, MA L Q. Discussion on aircraft metrology assurance guarantee system based on digital design and manufacturing collaboration [J]. Metrology and measurement technology, 2018, 38(2): 59-63. (in Chinese)

[4] 王博, 董登峰, 周维虎, 等. 面向激光跟踪仪跟踪恢复的合作目标视觉检测[J]. 光学精密工程, 2020, 28(2): 271-282.

- WANG B, DONG D F, ZHOU W H, et al. Visual detection of targetball for laser tracker target tracking recovery [J]. Optics and precision on engineering, 2020, 28(2): 271-282. (in Chinese)
- [5] 张滋黎, 郝继贵, 周虎, 等. 一种基于视觉引导的激光经纬仪自动测量系统[J]. 光电子激光, 2011, 22(1):115-119.
- ZHANG Z L, ZHU J G, ZHOU H, et al. Laser automatic theodolite measurement system based on vision guidance [J]. Journal of optoelectronics·laser, 2011, 22(1): 115-119. (in Chinese)
- [6] 崔艺涵. 大目标高精度位置姿态测量技术研究[D]. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2019.
- CUI Y H. Research on high precision position and attitude measurement technology of large target[D]. Changchun: University of Chinese Academy of Sciences (Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences), 2019. (in Chinese)
- [7] REN S, HE K, 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] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [J]. Computer vision and pattern recognition, 2016.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[J]. European conference on computer vision, 2016.
- [10] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// IEEE, 2017: 6517-6525.
- [11] 冯小雨, 梅卫, 胡大帅. 基于改进 Faster R-CNN 的空中目标检测[J]. 光学学报, 2018, 38(6): 250-258.
- FENG X Y, MEI W, HU D S. Aerial target detection based on improved Faster R-CNN [J]. Acta optica sinica, 2018, 38(6): 250-258. (in Chinese)
- [12] 王全东, 常天庆, 张雷, 等. 面向多尺度坦克装甲车辆目标检测的改进 Faster R-CNN 算法[J]. 计算机辅助设计与图形学学报, 2018, 30(12): 2278-2291.
- WANG Q D, CHANG T Q, ZHANG L, et al. An improved Faster R-CNN algorithm for detection of multi-scale tank armored vehicle targets [J]. Journal of computer-aided design and computer graphics, 2018, 30(12): 2278-2291. (in Chinese)
- [13] 程俊华, 曾国辉, 刘瑾. 基于深度学习的复杂背景图像分类方法研究[J]. 电子科技, 2020, 33(12): 62-69.
- CHENG J H, ZENG G H, LIU J. Research on complex background image classification method based on deep learning [J]. Electronic science and technology, 2020, 33(12): 59-66. (in Chinese)
- [14] ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[J]. 2017.
- [15] EVERINGHAM M, ESLAMI S, GOOL L V, et al. The pascal visual object classes challenge: a retrospective [J]. International journal of computer vision, 2015, 111(1): 98-136.

(本文编辑:刘圣晨)



第一作者:王国名(1986—),男,助理研究员,主要研究方向为光电精密测量技术及仪器。



通讯作者:周维虎(1962—),男,博士,中国科学院微电子所二级研究员,主要研究方向为光电精密测量技术与仪器、光电系统总体设计与集成等。



通讯作者:高豆豆(1994—),女,助理工程师,主要研究方向为机器视觉、深度学习和姿态测量。