

基于 YOLOv5s-T 和 RGB-D 相机的 螺栓检测与定位系统

王向周, 杨敏巍, 郑戌华, 梅云鹏

(北京理工大学 自动化学院, 北京 100081)

摘要: 机器人替代人工紧固角钢塔螺栓是解决高空作业安全问题的途径之一. 针对角钢塔螺栓紧固机器人的作业需求, 提出了一种基于神经网络和 RGB-D 相机的角钢塔主材螺栓检测与定位系统, 将轻量化的 YOLOv5s-T 网络应用于英特尔®实感™深度摄像头 D435i 采集的图像, 实现了角钢塔主材螺栓的实时检测、三维定位及重新排序等功能. 经实验验证, YOLOv5s-T 在基本不降低均值平均精度 (mean average precision, mAP) 的情况下, 推理速度提高约 31%; 用 RGB-D 相机测得的三维坐标计算相邻螺栓间距, 平均误差小于 1 mm. 对主材螺栓排序算法进行验证, RGB-D 相机正对螺栓组模板时, 模板的正确排序率不低于 95%, 可快速引导 6 自由度机械臂末端到达螺栓紧固点.

关键词: YOLOv5s-T; 螺栓检测; 三维定位; 排序

中图分类号: TP249

文献标志码: A

文章编号: 1001-0645(2022)11-1159-08

DOI: 10.15918/j.tbit.1001-0645.2021.339

Bolt Detection and Positioning System Based on YOLOv5s-T and RGB-D Camera

WANG Xiangzhou, YANG Minwei, ZHENG Shuhua, MEI Yunpeng

(School of Automation, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Replacing manual works with robots is a feasible solution for solving the safety problem of fastening bolts on the angle steel tower. In order to meet the operating requirements of the angle steel tower bolt fastening robot, a detecting and positioning system was proposed based on neural network and RGB-D camera for the main bolts of the angle steel tower. Applying the lightweight YOLOv5s-T network to the image of the Intel® RealSense™ depth camera D435i, the system was used to realize real-time detection, three-dimensional positioning and reordering the main bolts of the angle steel tower. Experiments show that YOLOv5s-T can improve the inference speed of the original algorithm by about 31% without reducing mAP (mean average precision) basically. Using three-dimensional coordinates measured by the RGB-D camera to calculate the distance between adjacent bolts, the average distance error is less than 1 mm. When the RGB-D camera is facing the bolt group template, the correct sorting rate of the template is above 95%. It can guide the end-effector of the 6-dof manipulator toward the target bolt within a short time.

Key words: YOLOv5s-T; bolt detection; three-dimensional positioning; ordering

角钢塔作为输电线路的重要组成部分, 主材螺栓的紧固性对于铁塔的防振动性能和结构稳定具有重要作用. 目前, 角钢塔塔身螺栓紧固多采用人工高空作业的方式, 效率不高且存在一定的安全隐患, 使

用自动化设备替代人工可以减少安全事故发生^[1]. 本文主要研究角钢塔主材螺栓的检测和粗定位, 快速引导螺栓紧固机器人的机械臂移动到待紧固螺栓上方, 以实现螺栓的精准定位. 目前, 常用的目标检

收稿日期: 2021-12-03

基金项目: 国家电网有限公司总部科技项目“角钢塔塔身螺栓紧固机器人研究”资助(5200-2020036147A-0-0-00)

作者简介: 王向周(1965—), 男, 博士, 教授, E-mail: wangxiangzhou@bit.edu.cn.

通信作者: 郑戌华(1976—), 女, 博士, 讲师, E-mail: zhengshuhua@bit.edu.cn.

测方法可以分为三类:图像处理、传统的机器学习和深度学习. SONG 等^[2]提出了一种基于形态学和小波变换的边缘检测算法来识别螺栓,经验证该算法能准确识别螺栓结构,但只能处理简单背景的图像. RAMANA 等^[3]使用传统的机器学习算法 Viola Jones 检测不同松动程度下的螺栓,准确率可达 97%,但图片中阴影、遮挡等情况会影响检测效果. 相比之下,深度学习克服了上述缺点,可以从复杂图像数据中学习高维特征,具有更强的鲁棒性. ZHANG 等^[4]为了改善传统 SSD 对小目标不敏感的缺点,提出了一种改进的基于多窗口多尺度融合的 SSD 网络, mAP 从 0.360 提高到 0.432. YANG 等^[5]采用改进的 YOLOv3-tiny 网络识别螺栓, mAP 从 0.813 增加到 0.839,但该方法只适用于相机垂直拍摄物体的角度. GE 等^[6]针对角钢塔螺栓在识别过程中易受遮挡的情况,用 ResNet50 代替 VGG16 骨干网络, mAP 从 0.896 提高到 0.924,但推理速度有限,识别一帧耗时 2.64 s. 上述方法主要集中在目标的二维定位上,采用机械臂紧固螺栓时,二维坐标难于实现引导作业,目前常用的方法是利用神经网络对二维图像进行目标检测,然后通过立体视觉技术进行三维定位. KANDA 等^[7]利用 YOLOv3 网络检测开关按钮,用 RGB-D 相机输出的点云估算按钮中心点的三维坐标,经验证当开关表面与 RGB-D 相机前表面平行,二者中心点相距 80 cm 时, x 、 y 、 z 方向的定位误差分别为 0.6、0.8 和 0.8 cm. HAN 等^[8]使用动态模板匹配算法根据二维强度图识别螺栓,达到 99.4% 的准确率, RGB-D 相机在

实验室条件下静态测量精度为 0.1 mm, 列车运行环境下动态测量精度为 0.5 mm. 但这些方法检测的都是单个目标,不需要确定多目标的顺序,不能满足螺栓组紧固要求.

本文提出一种基于神经网络和 RGB-D 相机的螺栓检测及定位系统,采用轻量化的 YOLOv5s-T 对螺栓进行识别并筛选主材螺栓,通过英特尔®实感™深度摄像头 D435i 对主材螺栓进行三维定位,同时将输出的坐标按照螺栓组紧固原则进行排序,实现了 6 自由度机械臂末端到螺栓紧固点的快速引导.

1 系统方法

1.1 难点分析

图 1 为某角钢塔塔身包括主材螺栓的 5 个不同段位,框内是规格为 M24 的主材螺栓,框外六角螺栓的规格为 M20. 如图 1 所示,角钢塔主材螺栓型号单一且分布集中,可据此分为不同的螺栓组模板,从左到右依次为模板 a、b、c、d、e,按照模板类型进行螺栓紧固. 要准确识别并定位主材螺栓,存在 3 个难点:神经网络具有庞大的参数量和计算量,搭载模型的硬件条件有限,故需要对网络模型进行轻量化操作;规格为 M24 的主材螺栓与其他规格的螺栓差别不大,神经网络检测时会检测出所有规格的螺栓,需要筛选出主材螺栓;紧固一组螺栓时需要按照一定的顺序分多次拧紧^[9],神经网络本身不具备目标排序的功能,故要对所有主材螺栓的三维坐标进行排序.

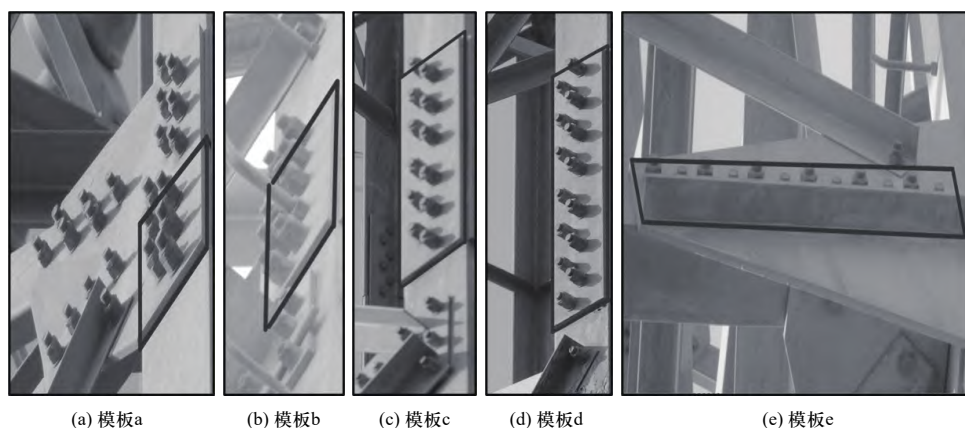


图 1 主材螺栓组

Fig. 1 Main material bolt set

1.2 螺栓检测定位系统

角钢塔螺栓检测及定位系统总体方案如图 2 所

示. 首先深度相机获取彩色图、深度图和点云,利用深度信息去除彩色图的背景,减少背景环境的干扰;

然后将处理过的彩色图送入 YOLOv5s-T 网络进行目标检测, 利用主材螺栓选取算法保留模板内主材螺栓的像素坐标; 最后结合点云获取三维坐标, 利用螺栓排序算法进行排序, 并将排序后的坐标保存到计算机。

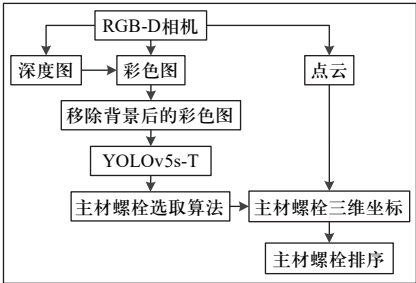


图 2 系统总方案
Fig. 2 General scheme of system

1.3 RGB-D 相机

RGB-D 相机选用了英特尔®实感™深度摄像头 D435i, 外形如图 3 所示, 包括立体视觉模块和颜色传感器两部分. 立体视觉模块由左、右成像器和红外投影仪组成, 同时搭载了英特尔实感视觉处理器, 使用主动红外立体技术来计算深度. 红外投影仪投射不可见的静态红外图案, 以提高低纹理场景的深度精度, 左、右成像器将数据发送到视觉处理器, 处理器将左、右图像相关联, 通过图像上点之间的移动来计算图像中每个像素的深度值. 整个系统的坐标系原点为颜色传感器的物理光心, x 轴、 y 轴和 z 轴的指向分别如图 3 所示. D435i 的各项技术规格如表 1 所示, 由表 1 可知, 相机可在户外使用, 使用范围满足机械臂长度要求, 采用 USB Type-C 接口, 体积小、视场大、分辨率高且易安装. 该相机采用了英特尔最新的深度感知硬件和软件, 集成度高, 英特尔官方网站提供了跨平台的开发软件 Intel RealSense SDK 2.0, 为二次开发提供了丰富的接口。

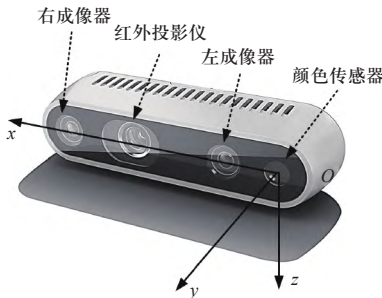


图 3 英特尔®实感™深度摄像头 D435i
Fig. 3 Intel® RealSense™ depth camera D435i

表 1 技术规格

Tab. 1 Technical specifications	
参数	规格
使用环境	室内/室外
使用范围/m	0.3 ~ 3.0
接口	USB Type-C 3.1 Gen 1
长×宽×高/(mm×mm×mm)	90 × 25 × 25
视场/(°)	69.4×42.5
分辨率/像素	1 920×1 080

2 实 现

2.1 YOLOv5s-T 网络

YOLOv5 网络模型共有 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 个版本, 不同的是网络的宽度与深度. 大的模型实时性差但精度高, 小的模型实时性好但精度差. YOLOv5s 虽为最小版本, 但该模型本身较大, 依旧不利于实际的工业场景应用. 本文基于 YOLOv5s 和常用的模型加速方法, 提出了轻量化网络模型 YOLOv5s-T. YOLOv5s-T 保留原网络 YOLOv5s 的整体结构, 如图 4 所示, 输入端用 Mosaic 数据增强方法, 主要思想是将 4 张图片进行随机裁剪, 再拼接到一张图上作为训练数据, 在丰富数据集的同时提升网络的训练速度, 降低模型的内存需求. 开头用 Focus 操作代替普通的下采样, 将每一个通道的图片进行切片, 特征图的尺寸减小为原来图片的 1/2, 而通道数变为原来的 4 倍, 其后加上一个 3×3 的点卷积来改变特征图的通道数, 极大地保留了原图的特征信息. 提取特征的骨干网络选用 CSPDarknet53, SPP 结构通过最大池化的方法进行多尺度特征融合. 头部网络在特征金字塔网络 (feature pyramid networks, FPN)+路径聚合网络 (path aggregation network, PAN) 的基础上, 用 CSPDarknet53 设计的 C3 结构代替普通的卷积网络, 加强网络特征融合能力. FPN 结构通过上采样将深层的语义信息传到底层, 补充浅层的语义信息, PAN 结构自底向上传达强定位特征, 从不同的主干层对不同的检测层进行特征聚合, 最后用 3 个检测头来进行多尺度的预测。

YOLOv5s-T 使用深度可分离卷积 (depthwise separable convolution) 和分组卷积 (group convolution) 结合通道混洗 (channel shuffle) 操作对网络进行轻量化, 如图 4 所示, 将 YOLOv5s 网络 Unit 和 Res Unit 中的 1×1 标准卷积修改为分组卷积, 3×3 标准卷积转换为

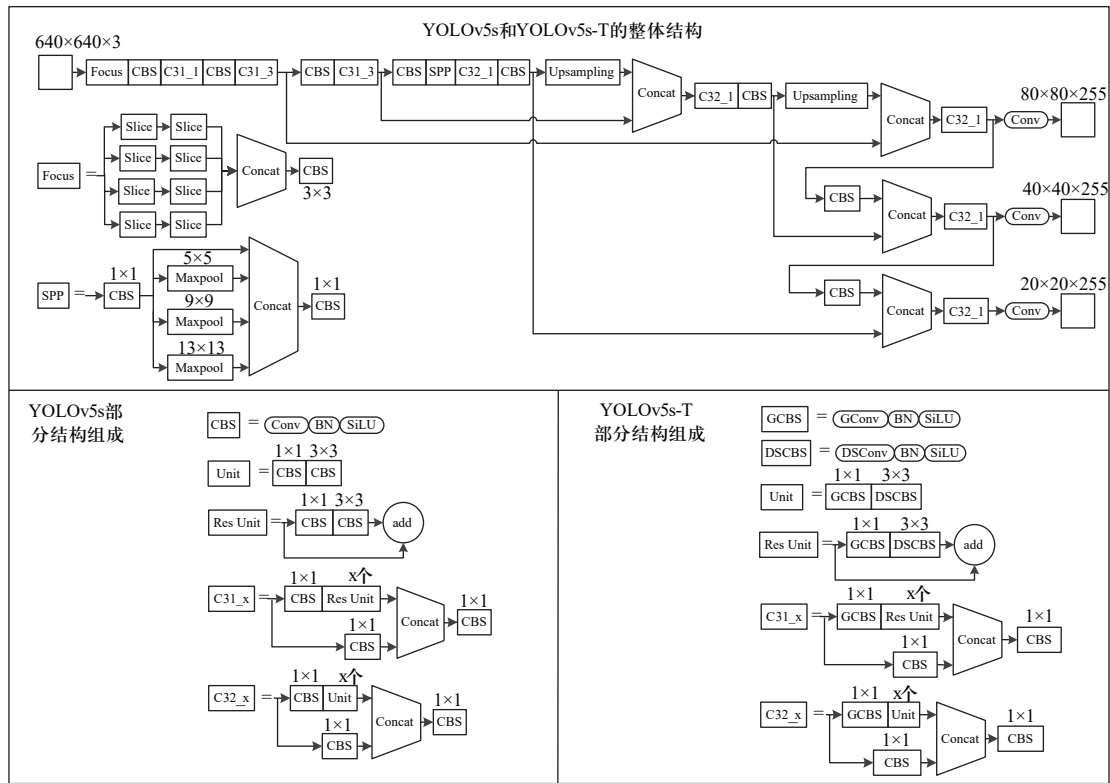


图 4 YOLOv5s 和 YOLOv5s-T 网络结构

Fig. 4 The network structures of YOLOv5s and YOLOv5s-T

深度可分离卷积, C31_x 和 C32_x 中第一个 1×1 标准卷积修改为分组卷积, 并在每个分组卷积后设置通道混洗操作以完成通道间的信息融合. 标准卷积是一种通道密集连接方式, 分组卷积是通道稀疏连接方式, 深度可分离卷积是特殊的分组卷积. 设输入特征图尺寸为 $C_i \times H \times W$, 输出特征图尺寸为 $C_o \times H \times W$, 卷积核尺寸为 $C_i \times k \times k$, 则标准卷积的参数量 (不考虑偏置) 为 $C_i \times k \times k \times C_o$, 计算量 (不考虑加法) 为 $C_i \times k \times k \times H \times W \times C_o$. 分组卷积是将输入特征图进行分组, 然后每组分别卷积, 最后用向量拼接 (concatenate) 的方式将输出数据组合起来, 如果将特征图分成 g 个组, 则每组输入特征图尺寸为 $\frac{C_i}{g} \times H \times W$, 对应的卷积核尺寸为 $\frac{C_i}{g} \times k \times k$, 每组输出特征图尺寸为 $\frac{C_o}{g} \times H \times W$, 参数量为 $C_i \times k \times k \times C_o \times \frac{1}{g}$, 计算量为 $C_i \times k \times k \times H \times W \times C_o \times \frac{1}{g}$, 二者均为标准卷积的 $\frac{1}{g}$. 分组卷积不仅可以极大地减少网络的参数量和计算量, 而且具有正则的效果, 防止过拟合. 通道混洗将每一个组的特征分散到不同的组后再进行下一次的卷积操作, 保证输出的特征图包含每一个组的特征, 促进各组之间的信息融合. 深度可分离卷积将标准卷积分

解成一个深度卷积和一个点卷积, 深度卷积将每个卷积核应用到每一个通道, 而点卷积用来组合通道卷积的输出^[10]. 深度卷积参数量为 $C_i \times k \times k \times 1$, 点卷积的参数量为 $C_i \times 1 \times 1 \times C_o$, 深度卷积计算量为 $C_i \times k \times k \times H \times W$, 点卷积的计算量为 $C_i \times 1 \times 1 \times H \times W \times C_o$, 故深度可分离卷积与标准卷积的参数量之比 r_p 和计算量之比 r_c 如式 (1) 所示. 深度可分离卷积虽然将卷积过程分为了两步, 但凭其轻量化的卷积方式, 大大减少了网络的参数量和计算量.

$$\begin{cases} r_p = \frac{C_i \times k \times k \times 1 + C_i \times 1 \times 1 \times C_o}{C_i \times k \times k \times C_o} = \frac{1}{C_o} + \frac{1}{k^2} \approx \frac{1}{k^2} \\ r_c = \frac{C_i \times k \times k \times H \times W + C_i \times 1 \times 1 \times H \times W \times C_o}{C_i \times k \times k \times H \times W \times C_o} = \frac{1}{C_o} + \frac{1}{k^2} \approx \frac{1}{k^2} \end{cases} \quad (1)$$

2.2 主材螺栓的选取

YOLOv5s-T 网络共检测 6 类, 分别为螺栓组模板 a、b、c、d、e 和螺栓, 检测完成后输出目标类别和目标检测框的左上角、右下角、中心点像素坐标. 设螺栓检测框中心点像素坐标为 (x_{pb}, y_{pb}) , 模板检测框的左上角和右下角像素坐标分别为 (x_{ptl}, y_{ptl}) 和 (x_{ptr}, y_{ptr}) , 满足式 (2) 的螺栓是主材螺栓, 统计模板内主材

螺栓的个数 n_c . 模板所含主材螺栓的真正个数 n_t 如表 2 所示, 当 $n_c = n_t$ 时, 保存模板类型和主材螺栓的像素坐标.

表 2 模板内主材螺栓的实际个数
Tab. 2 The actual number of bolts in the template

螺栓组模板类别	模板a	模板b	模板c	模板d	模板e
n_t	6	10	12	14	12

$$\begin{cases} x_{ptl} < x_{pb} < x_{ptr} \\ y_{ptl} < y_{pb} < y_{ptr} \end{cases} \quad (2)$$

YOLOv5s-T 检测完成后输出目标检测框上述 3 个点的像素坐标 (x_p, y_p) . 深度相机输出整张图片的点云 p , 即图片中每个像素点对应的三维坐标. 设图片的分辨率为 $w \times h$, 那么共有 $w \times h$ 个像素点. 设某一像素点为所有像素点中第 m 个, 则像素坐标到三维坐标的转换如式 (3)、式 (4) 所示, 将筛选后的主材螺栓检测框的左上角、右下角和中心点的像素坐标转换为三维坐标 (x, y, z) .

$$m = y_p w + x_p \quad (3)$$

$$\begin{cases} x = p[m][0] \\ y = p[m][1] \\ z = p[m][2] \end{cases} \quad (4)$$

2.3 主材螺栓排序

在高强度螺栓连接中, 弹性相互作用影响较大, 螺栓的预紧力也会有大幅度降低, 为了减轻这种影响, 螺栓组要按照一定的顺序分多次拧紧其中的螺栓. 分次拧紧螺栓时只改变扳手的扭矩, 不改变螺栓组中螺栓的紧固顺序. YOLOv5s-T 网络按照置信度高低的顺序输出坐标信息, 不符合螺栓组的紧固规则, 需要对模板内的主材螺栓重新排序. 5 个模板主材螺栓的分布均呈平行四边形, 如图 5 所示, 按照先

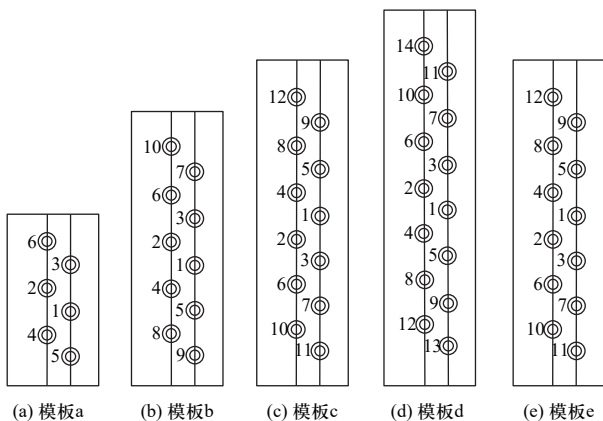


图 5 螺栓组模板紧固顺序

Fig. 5 Fastening sequence of bolt group template

中间后两边、对角就近的原则对螺栓组进行紧固, 图片中螺栓上方的数字为模板螺栓组的紧固顺序.

主材螺栓排序前需要确定螺栓组模板, 不同模板排序算法实现不同但理念相同, 即根据主材螺栓间的距离关系重新排序. 由图 5 可知, 模板 c 和模板 e 主材螺栓紧固顺序相同, 所以可归为一类. 以模板 a 为例, 先找出最大距离对应的两颗螺栓, 由于所有模板内主材螺栓分布均呈中心对称状, 故任选其中一颗作为第一颗螺栓记为 6 号, 然后从剩下的螺栓中找到距离 6 号螺栓最近的一颗螺栓, 记为 3 号; 再从剩下的螺栓中找到距离 3 号螺栓最近的一颗螺栓, 记为 2 号, 以此类推, 直到所有螺栓都被编号, 其他模板排序算法理念与之相同. 根据螺栓序号保存主材螺栓的三维坐标.

3 实 验

3.1 数据集

受到实验室条件和角钢塔螺栓组模板数据量的限制, 用自制的 5 个螺栓纸板模拟螺栓组模板验证系统的各项性能. 如图 6 所示, 从左到右依次为模板 a、b、c、d、e. 数据集是从不同角度、不同距离和不同背景下采集的 5 个模板图片, 共 1 150 张, 通过调整亮度和仿射变换的方法进行数据增强, 将数据集扩充到 4 600 张. 利用图像标注工具 Labellmg 进行人工标注, 并按照 6 : 2 : 2 的比例将图片分为训练集、验证集和测试集.

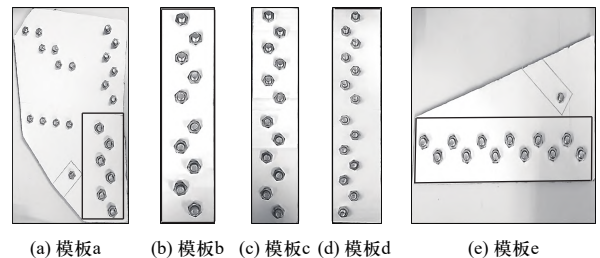


图 6 模拟螺栓组模板

Fig. 6 Template of simulated bolt set

3.2 实验细节

视觉采集系统由英特尔®实感™深度摄像头 D435i 及其自带的三脚架、USB 线和自制的螺栓纸板组成. 数据在服务器上训练, 硬件配置为带两块显卡的 LT-6 028 GPU 服务器, 其中显卡型号为 GeForce GTX 1080 Ti, 内存均为 11 G, 操作系统为 Ubuntu 16.04. 模板和螺栓的检测、定位在电脑端实现, 电脑配置为 Intel(R) Core(TM) i5-10210Y CPU. 系统运行所配

置的软件为 python3.7.9, torch 1.8.0+cpu. 实验用到了 OpenCV 图形开发库和系统基本依赖以及 RGB-D 相机的 API 接口 Intel Realsense SDK 2.0 和标定软件 Intel Realsense CalibrationTool API 2.11.0.0 等.

为了加快模型收敛, 网络在训练时使用官方提供的预训练权重文件初始化部分网络权重参数. 使用 SGD 优化器进行参数更新, 余弦退火方式进行学习率衰减, 学习率初值设置为 0.01, 动量为 0.937, 训练的类别数量设置为 6, 迭代次数为 60, 批数量为 32, 图片分辨率设置为 640×480.

3.3 实验结果与分析

3.3.1 模型性能

实验采用 mAP@0.5、mAP@0.5:0.95 两个评价指标来评估网络模型的检测性能, 用训练参数量 T 、浮点运算数 F 和摄像头实时检测速率 v 来评估模型的复杂度.

表 3 为 YOLOv5s-T、YOLOv5s、YOLOv3 和 SSD 的检测结果, 相对于原来的 YOLOv5s 模型, 本文提出的 YOLOv5s-T 的 mAP@0.5 降低 0.02, mAP@0.5 : 0.95 降低 0.03, T 减少 36.2%, F 减少 42.1%, v 提高 31%; 与 YOLOv3 相比, YOLOv5s-T 的 mAP@0.5 提高 0.20, mAP@0.5 : 0.95 提高 0.23, T 减少 92.7%, F 减少 71.0%, v 提高 445%; 与 SSD 相比, YOLOv5s-T 的 mAP@0.5 提高 0.27, mAP@0.5 : 0.95 提高 0.30, T 减少 81.5%, F 减少 69.1%, v 提高 342%. 综合考虑模型的准确度和复杂度, 选择轻量化模型 YOLOv5s-T.

表 3 检测结果

Tab. 3 Detection result

算法	mAP@0.5	mAP@0.5 : 0.95	T /Million	F /G	v /FPS
YOLOv5s-T	0.94	0.74	4.52	9.50	5.83
YOLOv5s	0.96	0.77	7.08	16.40	4.45
YOLOv3	0.74	0.51	61.50	32.80	1.07
SSD	0.67	0.44	24.41	30.74	1.32

3.3.2 定位性能

RGB-D 相机在运动过程中会导致螺栓的三维坐标发生变化, 但螺栓间距是固定的. 设相邻两颗螺栓的三维坐标分别为 (x_p, y_p, z_p) 和 (x_q, y_q, z_q) , 相邻螺栓间的距离

$$d = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2 + (z_p - z_q)^2} \quad (5)$$

设测量距离为 d_m , 实际距离为 d_t , 则距离误差

$$e = |d_m - d_t| \quad (6)$$

通过 e 对系统的定位性能进行分析, 以模板 b 为例, 螺栓用排序之后的序号表示, 如图 5(b) 所示, 计算相邻螺栓间的距离. 为了保证数据的可靠性, 每组实验重复 10 次, 取平均值. 结果如表 4 所示, 忽略螺栓模板制造工艺的误差, 相邻螺栓间距离的最大误差不超过 3 mm, 平均误差不超过 1 mm.

表 4 定位误差

Tab. 4 Location error

序号-序号	d_m /mm	d_t /mm	e /mm
9-8	32.5	34.3	1.8
8-5	32.5	34.7	2.2
5-4	32.5	31.4	1.1
4-1	32.5	33.4	0.9
1-2	43.9	44.3	0.4
2-3	32.5	34.2	1.7
3-6	32.5	31.1	1.4
6-7	32.5	33.6	1.1
7-10	32.5	34.8	2.3
平均值	33.8	34.6	0.8

3.3.3 主材螺栓排序

螺栓组模板固定不动, 将 RGB-D 相机置于螺栓组模板的不同方位对主材螺栓进行排序. 以模板 b 为例, 图 7 为 RGB-D 相机正对螺栓组模板时的三视图, 设 RGB-D 相机前表面中心点为 O_c , 螺栓组模板前表面中心点为 O_t , 两点间的距离 l 为 0.7 m. 实验分为 3 组, 固定螺栓组模板不动, 分别在 x - O - y 、 y - O - z 、 z - O - x 平面内旋转 RGB-D 相机, 验证 RGB-D 相机在不同角度下拍摄螺栓组模板时主材螺栓排序算法的性能. 第 1 组以 RGB-D 相机正对螺栓组模板的 x 轴为基准线, O_c 为旋转中心, 将 RGB-D 相机的机身在 x - O - y 平面旋转一定的角度, 如图 7(a) 所示, 旋转后的 x 轴与基准线间的夹角为 α ; 第 2 组以 RGB-D 相机正对螺栓组模板的 z 轴为基准线, O_t 为旋转中心, l 为半径, 将 RGB-D 相机在 y - O - z 平面旋转一定的角度, 如图 7(b) 所示, 旋转后的 z 轴与基准线之间的夹角为 β ; 第 3 组以 RGB-D 相机正对螺栓组模板的 z 轴为基准线, O_t 为旋转中心, l 为半径, 将 RGB-D 相机在 z - O - x 平面旋转一定的角度, 如图 7(c) 所示, 旋转后的 z 轴与基准线之间的夹角为 γ . 其他螺栓组模板的实验设计与此相同, 对 5 个螺栓组模板的主材螺栓进行排序, 每组实验重复 n_s 次, 统计每个螺栓组模板正确排序的次数 n_t , 则螺栓组模板的正确排序率

$$\eta = \frac{n_r}{n_s} \times 100\%$$

(7)

RGB-D 相机在各个平面内旋转后螺栓组模板的正确排序率如表 5~表 7 所示, η_a 、 η_b 、 η_c 、 η_d 、 η_e 分别表示模板 a、b、c、d、e 的正确排序率, 当 RGB-D 相机正对着螺栓组模板, 即 α 、 β 、 γ 都为 0° 时, 每个螺栓组模板的正确排序率均在 95% 以上. 由于神经网络存在一定的误检, 螺栓检测不准确导致无法确认螺栓组模板类别. 此外, 深度相机还受物体颜色和表面光滑度的影响, 深色物体对红外光的吸收能力非

常强, 导致投射在深色物体上的部分红外光无法返回, 左、右成像器因为无法接收到返回的红外光, 也就无法测量物体的深度. 漫反射是物体可以清晰成像的基础, 当物体表面过于光滑时发生镜面反射, 此时只有当被测物体位于深度相机的发射端和接收端的中垂线时, 才能接收到反射光线, 故当物体表面超过一定的光滑度时, 深度相机测量精度会急剧下降, 甚至无法测量深度. 由于上述原因, 实验的 100 帧图片中存在极个别螺栓组模板无法被正确排序的情况.

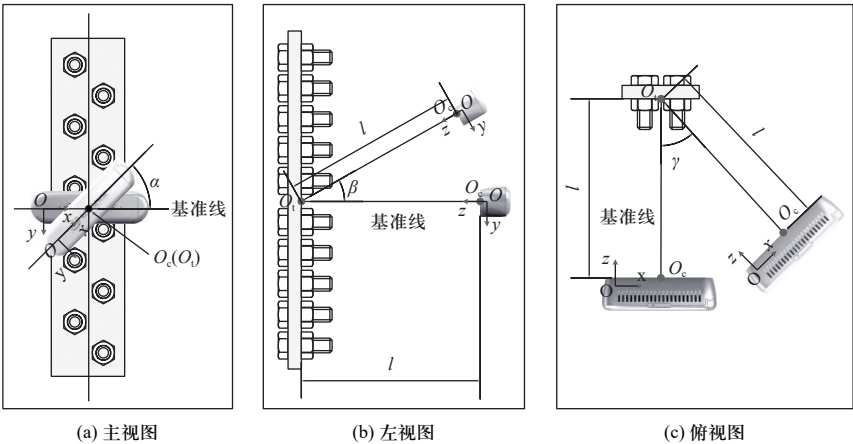


图 7 模板 b 主材螺栓排序三视图

Fig. 7 Three views of sorting of main material bolts of template b

表 5 x - O - y 平面模板正确排序率

Tab. 5 The correct sorting rate of templates in the x - O - y plane							
$\alpha/(^\circ)$	$\beta/(^\circ)$	$\gamma/(^\circ)$	$\eta_a/\%$	$\eta_b/\%$	$\eta_c/\%$	$\eta_d/\%$	$\eta_e/\%$
0	0	0	97	98	97	96	98
30	0	0	95	96	95	95	0
60	0	0	0	99	98	96	97
-30	0	0	0	95	93	95	96
-60	0	0	96	98	97	97	0

表 6 y - O - z 平面模板正确排序率

Tab. 6 The correct sorting rate of templates in the y - O - z plane							
$\alpha/(^\circ)$	$\beta/(^\circ)$	$\gamma/(^\circ)$	$\eta_a/\%$	$\eta_b/\%$	$\eta_c/\%$	$\eta_d/\%$	$\eta_e/\%$
0	0	0	97	98	97	96	98
0	30	0	96	96	95	95	95
0	60	0	97	96	99	97	97
0	-30	0	96	98	95	96	98
0	-60	0	97	95	97	96	97

如表 5 所示, 当 RGB-D 相机在 x - O - y 平面旋转时, 模板 b、c、d 不受旋转角度的影响, 正确排序率均在

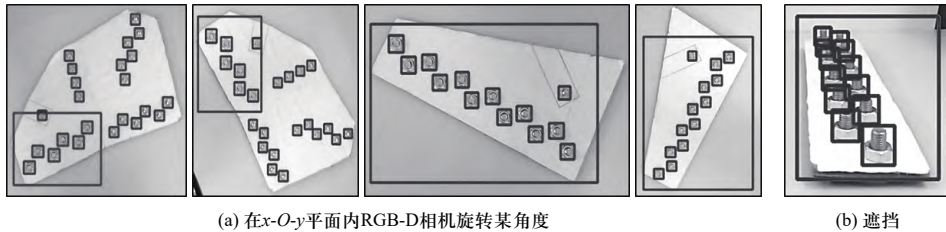
表 7 z - O - x 平面模板正确排序率

Tab. 7 The correct sorting rate of templates in the z - O - x plane							
$\alpha/(^\circ)$	$\beta/(^\circ)$	$\gamma/(^\circ)$	$\eta_a/\%$	$\eta_b/\%$	$\eta_c/\%$	$\eta_d/\%$	$\eta_e/\%$
0	0	0	97	98	97	96	98
0	0	30	96	96	95	95	98
0	0	60	98	95	96	97	96
0	0	-30	96	97	95	98	97
0	0	-60	97	97	97	96	98

95% 以上; 当 $\alpha=60^\circ$ 或者 $\alpha=-30^\circ$ 时, 模板 a 的正确排序率为 0; 当 $\alpha=30^\circ$ 或者 $\alpha=-60^\circ$ 时, 模板 e 的正确排序率为 0. 如图 8(a) 所示, 受到目标检测框形状的限制, 其他规格的螺栓在旋转 RGB-D 相机后也被包含在模板内, 检测的主材螺栓个数会大于实际模板主材螺栓个数, 无法确定模板类型, 进而不能执行下一步排序操作. 如表 6、表 7 所示, 当 RGB-D 相机在 y - O - z 平面旋转, 即 $-60^\circ < \beta < 60^\circ$ 时, 5 个螺栓组模板的正确排序率均在 95% 以上. 当 RGB-D 相机在 z - O - x 平面旋转, 即 $-60^\circ < \gamma < 60^\circ$ 时, 5 个螺栓组模板的正确排序率均在 95% 以上. 当 RGB-D 相机在 y - O - z 和 z - O - x 平

面内旋转时,随着旋转角度绝对值的进一步增大,图片中的螺栓开始相互遮挡,螺栓检测框出现重叠现象

象,如图 8(b) 所示,相邻螺栓间的距离测量产生较大误差,影响螺栓组模板正确排序。



(a) 在 $x-O-y$ 平面内RGB-D相机旋转某角度

(b) 遮挡

图 8 螺栓组模板无法被正确排序的两类情况

Fig. 8 Two situations when bolt group templates cannot be sorted correctly

4 结 论

针对角钢塔主材螺栓检测与定位问题,结合轻量化的YOLOv5s-T网络和英特尔®实感™深度摄像头D435i,提出了角钢塔主材螺栓的检测与定位系统.通过一系列对比试验,验证了轻量化网络模型YOLOv5s-T在基本不降低mAP的情况下,推理速度提高31%.通过比较相邻螺栓的测量距离与真实距离,验证系统的平均误差不超过1 mm.对主材螺栓排序算法进行验证,发现当RGB-D相机正对螺栓组模板时,螺栓组模板的正确排序率不低于95%.在 $x-O-y$ 平面旋转RGB-D相机时,螺栓组模板的正确排序率会受到模板类别和旋转角度的影响.在 $y-O-z$ 和 $z-O-x$ 平面旋转RGB-D相机时,当旋转角度绝对值小于 60° ,螺栓组模板的正确排序率在95%以上。

角钢塔主材螺栓的检测与定位系统实现了角钢塔主材螺栓的实时检测、三维定位和重新排序等功能,为快速导引6自由度机械臂末端到达螺栓紧固点进行精确定位的二次成像及紧固作业奠定了基础。

参考文献:

- [1] 宋屹峰,王洪光,李贞辉,等.基于视觉方法的输电线断股检测与机器人行为规划[J].机器人,2015(2):204-211+223.
SONG Yifeng, WANG Hongguang, LI Zhenhui, et al. Vision based transmission line broken strand detection and robot behaviour planning[J]. Robot, 2015(2): 204-211+223. (in Chinese)
- [2] SONG Q, XIAO X, JIANG H. The automatic recognition of large ball valve sealing bolt based on digital image[J]. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015, 8(6): 311-320.
- [3] RAMANA L, CHOI W, CHA Y J. Fully automated vision-based loosened bolt detection using the Viola-Jones algorithm[J]. Structural Health Monitoring, 2019, 18(2): 422-

434.

- [4] ZHANG J, SU Z, XING Z. An improved SSD and its application in train bolt detection[C]//Proceedings of the 4th International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2019. Singapore: Springer, 2019: 97-104.
- [5] YANG J, XIN L, HUANG H, et al. An improved algorithm for the detection of fastening targets based on machine vision[J]. Computer Modeling in Engineering & Sciences, 2021, 128(2): 779-802.
- [6] GE L, XU J, WU X, et al. Research on tower bolt identification technology based on convolution network[C]//Proceedings of 2020 International Conference on Artificial Intelligence, Computer Networks and Communications (AICNC). Yunnan, China: [s. n.], 2020: 27-30.
- [7] KANDA T, MIYAKAWA K, HAYASHI J, et al. Locating mechanical switches using RGB-D sensor mounted on a disaster response robot[C]//Proceedings of 2020 Intelligent Robotics and Industrial Applications Using Computer Vision Conference, IRIACV 2020. Burlingame, USA: [s. n.], 2020.
- [8] HAN Q, WANG S, FANG Y, et al. A rail fastener tightness detection approach using multi-source visual sensor[J]. Sensors, 2020, 20(5): 1367.
- [9] 张博超. 航发低压涡轮轴盘连接预紧力精确控制研究[D].大连:大连理工大学,2018.
ZHANG Bochao. Research on precision control of pretightening force of aeroengine low pressure rotor[D]. Dalian: Dalian University of Technology, 2018. (in Chinese)
- [10] 周治国,刘开元,郑翼鹏,等.一种基于深度学习的高速无人艇视觉检测实时算法[J].北京理工大学学报,2021,41(7): 758-764.
ZHOU Zhiguo, LIU Kaiyuan, ZHENG Yipeng, et al. A real-time algorithm for visual detection of high-speed unmanned surface vehicle based on deep learning[J]. Transactions of Beijing Institute of Technology, 2021, 41(7): 758-764. (in Chinese)

(责任编辑:李兵)