

文章编号:1001-5078(2024)12-1864-07

· 激光应用技术 ·

# 激光切割机器人视觉图像目标标注研究

熊艳飞, 刘登邦

(江西应用科技学院 智能制造工程学院, 江西 南昌 330100)

**摘要:**在激光切割的工业环境下会存在大量干扰元素,例如电磁干扰、振动、烟尘和颗粒物、外部环境以及光源等,当前的结合类别学习方法,对目标抽象图形中凸显区域缺少单独目标对齐过程,导致目标特征关联性不强,标注结果不准。提出基于改进YOLOv5s的激光切割机器人视觉图像目标标注方法。利用输入端、池化层、共享全连接层等搭建改进YOLOv5s模型,该网络使用最大池化与平均池化生成两幅激光切割机器人视觉图像,根据通道维度连接图像特征实现激光视觉图像目标粗定位,结合调制因子和目标检测损失实现目标特征对齐。在目标特征对齐后确定激光切割机器人视觉图像关键凸显区域帧,通过对激光切割机器人历史标注图像实施半监督训练,确定图像空间区域关联,根据区域关联进行激光视觉图像目标标注。实验结果表明:所提方法的激光切割机器人视觉图像目标标注的交并比与准确率高、速度快,拥有极强的鲁棒性。

**关键词:**改进YOLOv5s;激光切割机器人;视觉图像;目标标注;目标定位

**中图分类号:**TN249;TN957.52 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2024.12.009

## Research on visual image object annotation of laser cutting robot

XIONG Yan-fei, LIU Deng-bang

(School of Intelligent Manufacturing Engineering, Jiangxi Institute of Applied Science and Technology, Nanchang 330100, China)

**Abstract:**In the industrial environment of laser cutting, there are a large number of complex backgrounds and other interfering elements such as equipment. The current combination of category learning methods lacks a separate target alignment process for the highlighted areas in the abstract target graphics, resulting in weak correlation between target features and inaccurate annotation results. Propose a visual image object annotation method for laser cutting robots based on improved YOLOv5s. By utilizing input terminals, pooling layers, and shared fully connected layers, an improved YOLOv5s model is built. This network uses max pooling and average pooling to generate two laser cutting robot visual images. Based on the channel dimension, the image features are connected to achieve rough target localization in the laser visual image, and target feature alignment is achieved by combining modulation factors and target detection losses. After aligning the target features, the key highlighted area frames of the laser cutting robot's visual image are determined. By implementing semi supervised training on the historical annotated images of the laser cutting robot, the spatial region associations of the images are determined, and laser visual image target annotation is performed based on the region associations. The experimental results show that the proposed method for laser cutting robot visual image

基金项目:2022年江西省科技厅科技课题项目(No. GJJ2203317)资助。

作者简介:熊艳飞(1988-),女,硕士,讲师,研究方向为自动化控制。E-mail:xiongYanfei0631Aa@163.com

收稿日期:2024-08-30

target annotation has high intersection to union ratio, high accuracy, fast speed, and strong robustness.

**Keywords:** improve YOLOv5s; laser cutting robot; visual images; target annotation; target location

## 1 引言

激光切割机器人<sup>[1]</sup>是一种集成了激光切割技术和机器人技术的先进设备,其主要工作原理是激光器发射出的激光,经过镜面传输,再利用聚光反射镜照射在被加工物体上,从而对被加工物体表面施加热量,使其温度迅速升高,在高温作用下该点快速熔化或汽化,此过程中需要保证加工点轨迹与激光头运动轨迹相匹配,从而实现精准切割的目标<sup>[2]</sup>。而当前,如何快速识别和标记被加工物体,明确目标对象在图像中的位置<sup>[3]</sup>,以实现高精度机械加工,是激光切割机器人应用领域亟待解决的核心问题。

图像目标标注是一种数据准备方法,而目标检测是一种应用这些标注数据来实现自动识别和对象定位的技术,两者是紧密相关的。但是在实际中针对激光切割机器人目标标注的研究极少,当前诸多研究主要是将机器人目标检测领域的研究迁移到这一领域中来,例如王薇<sup>[4]</sup>等人提出基于 CNN 的机器人视觉光谱图像目标识别技术。但是其所用的 Hessian 矩阵的计算涉及二阶导数,计算量相对较大,对于大规模图像处理任务会面临计算资源不足的问题,导致检测质量下降。黄怡蒙<sup>[5]</sup>等人提出了融合深度学习的机器人目标检测与定位方法。这一方法对 Tiny-YOLOV3 进行了改进,但由于其基础网络结构较为简单,卷积层数量较少,相比于更复杂的模型其特征提取能力仍存在不足,导致在目标遮挡、形变情况下的检测精度下降。郭雨青<sup>[6]</sup>提出了图像增强水下自主机器人目标识别方法,该方法所使用的对比度受限的自适应直方图均衡化技术能够在一定程度上提升图像的视觉质量,但它会使图像的某些细节丢失,导致目标检测质量下降。吴建华<sup>[7]</sup>提出了多源信息融合的巡检机器人目标检测算法。该方法所使用的深度可分离卷积层虽然能够减少计算量和参数量,但无法充分提取目标的所有关键特征,特别是对于小型目标或具有复杂纹理的目标,深度可分离卷积可能无法捕捉到足够的细节信息,导致目标检测精度下降。

在激光切割的工业环境下,通常存在大量的电磁干扰、振动、烟尘和颗粒物、外部环境以及光源等干扰元素,这些因素对目标检测和标注造成困难,为

此提出一种基于改进 YOLOv5s 的激光切割机器人视觉图像目标标注方法。

## 2 改进 YOLOv5s 下激光切割视觉图像目标粗定位

改进 YOLOv5s 模型结构如图 1 所示。

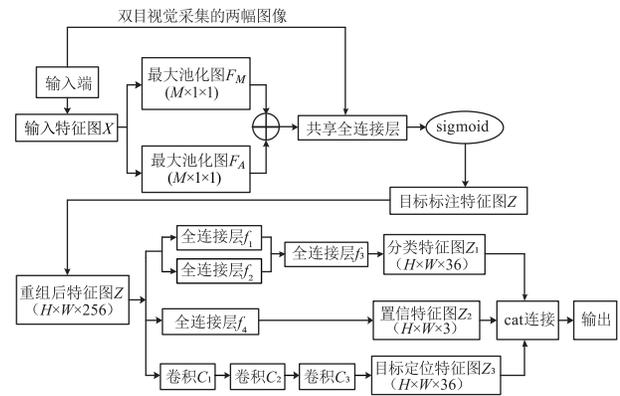


图 1 改进 YOLOv5s 模型示意图

Fig. 1 Schematic diagram of improved YOLOv5s model

在改进 YOLOv5s 中<sup>[8-9]</sup>,利用注意力机制降低模型传递特征数据中的干扰,突出图像中的关键特征,抑制不重要的背景信息,从而提高特征提取的质量。分析图 1 可知,在改进 YOLOv5s 模型中输入特征图  $X$ ,通过最大池化与平均池化操作,生成两张  $M \times 1 \times 1$  的特征图。设定最大池化图为  $F_M$ ,平均池化图是  $F_A$ ,将激光切割机器人通过双目视觉采集的两幅图像,依次引入共享全连接层,运用 Sigmoid 激活函数增强通道权重,图像目标生成过程为:

$$A_c(X) = S(M_{LP}(D_p(X); N_p(X))) \\ = S(U_1(U_0(F_A)) + U_1(U_0(F_M))) \quad (1)$$

其中,  $S$  为 Sigmoid 激活函数;  $D_p$  为平均池化操作;  $N_p$  为最大池化操作;  $M_{LP}$  为共享全连接层;  $U_0$ 、 $U_1$  依次为共享全连接层的第 1 层与第 2 层。

在对切割目标图像的分类任务中,对目标特征图  $Z$  采取重组操作,令其成为  $H \times W \times 256$  的特征图。把获得的特征图代入两个全连接层  $f_1$ 、 $f_2$ ,再整合输出结果,代入全连接层  $f_3$ ,得到  $H \times W \times 36$  的分类特征图,命名为  $Z_1$ 。

在置信度计算过程中,将其拟作一个软标签分类任务,重组操作切割目标特征图  $Z$ ,代入全连接层  $f_4$  对置信度特征实施整合运算,得到一幅  $H \times W \times 3$  的置信度特征图,命名为  $Z_2$ 。

在定位任务中,采用卷积  $C_1$  与卷积  $C_2$  分别对

定位特征采取整合运算,再使用卷积层  $C_3$  获得目标定位特征图  $Z_3$ 。最终是把三个任务的输出结果依照通道维度进行 cat 连接操作,结果记作:

$$Z' = \text{cat}(Z_3, Z_2, Z_1) \quad (2)$$

其中,  $\text{cat}(\cdot)$  为连接函数,该函数的作用是将三个任务的输出结果沿着指定的维度连接起来,形成一个新的特征图。在目标粗定位阶段改进 YOLOv5s 会预先定义一组锚框,这些锚框是一些具有不同尺寸和宽高比的矩形框,它们分布在特征图的每个位置上,用于匹配不同大小和形状的目标。改进 YOLOv5s 会预测每个锚框的调整参数生成新的边界框,这些边界框是对锚框的修正,以更好地匹配图像中的实际目标,从而实现目标粗定位。

### 3 激光切割机器人视觉图像目标标注

#### 3.1 切割机器人标注目标特征对齐

为衡量源域样本与目标域样本的特征差异,采用改进 YOLOv5s 粗定位后,融合多类型的损失函数对齐激光切割机器人视觉图像目标,提升标注效率。过程如下:

设定切割目标特征提取器是  $E$ ,切割目标图像域分类器是  $D_g$ ,则切割目标图像检测损失为:

$$H_F = -\frac{1}{b_s} \left[ Z' y_i \log(D_g(E(x_i))) + (1 - y_i) \log(1 - D_g(E(x_i))) \right] \quad (3)$$

其中,  $i$  为样本索引值;  $y_i$  表示真实的样本标签;  $b_s$  为训练样本个数。

为避免忽略易分类样本,采用焦点损失算法,在交叉熵中引入调制因子,具体过程为:

$$F_L(q_i) = -H_F f(q_i) \ln(q_i) \quad (4)$$

其中,

$$f(q_i) = (1 - q_i)^\beta \quad (5)$$

其中,  $\beta$  为控制权重因子;  $q_i$  为图像被划分成目标域的概率;  $f(q_i)$  为目标分类概率函数。设定训练样本分类损失为  $L_g$ ,测试样本分类损失为  $L_b$ ,则特征对齐的总损失的计算公式为:

$$L_k = \frac{F_L(q_i)(L_g + L_b)}{2} \quad (6)$$

根据源域和目标域的样本特征,对于切割机器人标注目标特征进行对齐处理,具体的计算公式如下:

$$W = \frac{\min L_k(MMD(E(X_s), E(X_t)))}{\lambda} \quad (7)$$

其中,  $M$  表示特征对其函数;  $E(X_s)$ 、 $E(X_t)$  分别表示源域与目标域样本特征;  $\lambda$  表示一个超参数,用于平衡源域与目标域样本的重要性。

#### 3.2 基于关键帧选择的视觉图像目标标注

当待切割目标轨迹断裂产生检测误差时,待标注目标位置的像素值和上一帧正确的目标像素值之间的差异性较大。将这一属性拟作评估关键帧的核心点,评估条件为:

$$V_t < \delta_2 V_{t-1}, V_t > \delta_1 V_{t-1} \quad (8)$$

其中,  $V_t$  为第  $t$  帧待标注目标像素;  $\delta_1$ 、 $\delta_2$  均为常数;  $V_{t-1}$  为上一帧正确的目标像素。

由于正常工况下目标在视觉图像内的运动为持续且规律的,将某帧图像前后标注距离和预期值差距较高的图像帧拟作关键帧,评估条件为:

$$v_t > \frac{\zeta}{t} \sum_{i=1}^t v_i \quad (9)$$

其中,  $v_t$  为理想标注距离(单位:mm);  $v_i$  为第  $i$  张图像目标与前一帧图像目标的标注间距(单位:mm);  $\zeta$  为常数。倘若当前帧的  $v_t$  超出历史均值,则该帧图像  $V_t$  即为关键帧。

获得关键帧后,完成待标注目标的分析与识别,搜集激光切割机器人历史标注图像,对未标注图像与历史标注图像实施半监督训练,了解待标注图像内广泛的空间区域关联。半监督训练损失函数描述为:

$$I_G = V_t I_{ce} + \eta_{adv} I_{adv} + \eta_{semi} I_{semi} \quad (10)$$

式中,  $\eta_{adv}$ 、 $\eta_{semi}$  均为损失权重;  $I_{adv}$ 、 $I_{semi}$  依次为对抗损失与二次损失;  $I_{ce}$  为标准二次损失。

对于激光视觉特征图中两个像素点之间的区域关联性进行计算,具体的公式如下:

$$O_{ji} = \frac{\exp(Q_i, K_j)}{\sum_{i=1}^{n_p} \exp(Q_i, K_j)} \quad (11)$$

其中,  $Q_i$ 、 $K_j$  依次为激光视觉特征图  $Q$  第  $i$  个方位的像素与激光视觉特征图  $K$  第  $j$  个方位像素相对的矢量;  $n_p$  为迭代次数。

在图1的注意力机制板块中代入加权特征图,则展示的图像性质为:

$$W_j = \vartheta \sum_{i=1}^{n_p} (O_{ji} + I_c T_i) + X_j \quad (12)$$

其中,  $T_i$  为特征图  $T$  第  $i$  个位置像素的对应矢量;  $X_j$  为特征图  $X$  第  $j$  个像素的对应矢量;  $\vartheta$  为比例参数。

根据区域关联归一化处理激光视觉图像矩阵,

输出图像目标标注结果,记作:

$$\bar{G}_{SN}(G) = \frac{W_j L_k G}{\theta(G)} \quad (13)$$

其中,  $\theta(G)$  为激光视觉图像参数集  $G$  的二范数。

#### 4 实验分析

为表明所提方法可靠性,对其进行实验分析,实验平台为工业化软件平台。使用如图2所示的6自由度工业激光切割机器人,其体积小,加工速度可达5 m/min,系统耗电低于6 kW/h,可搭载离线编程软件,重复定位精度为  $\pm 0.06$  mm。



图2 激光切割机器人示意图

Fig. 2 Schematic diagram of laser cutting robot

利用视觉传感器收集包含激光切割目标(如金属板、零件等)的图像数据集,以确保模型的泛化能力。实验数据集示例如图3所示。



图3 实验数据集示例

Fig. 3 Example of experimental dataset

使用 TensorBoard 来跟踪训练过程中的损失函数值,以此检验在目标检测过程中的未添加和添加注意力机制的 YOLOv5s 模型性能,具体结果如图4所示。

分析图4中的结果可知,随着训练轮数的增加,添加注意力机制的 YOLOv5s 模型在迭代次数90次后曲线趋于稳定,而未添加注意力机制的 YOLOv5s 模型在迭代次数155次后曲线趋于稳定,由此说明添加注意力机制的 YOLOv5s 模型更优,可以保证目标检测精度。

以汽车金属板切割为例,采用所提方法进行图像标注过程中,具体的实现过程如图5所示。

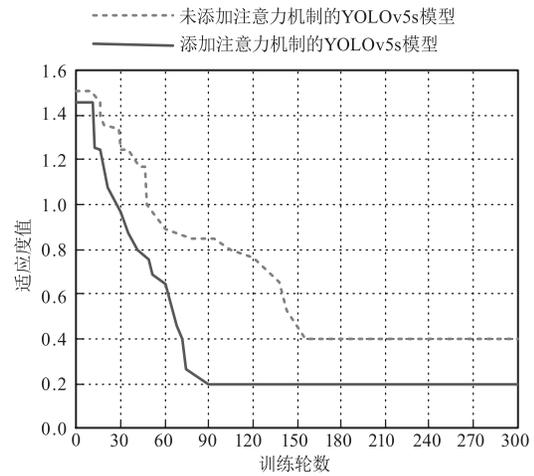


图4 YOLOv5s 模型性能

Fig. 4 Performance of YOLOv5s model



(a) 视觉图像目标特征点提取



(b) 视觉图像目标粗定位



(c) 视觉图像目标特征对齐



(d) 视觉图像目标标注

图5 图像标注结果

Fig. 5 Image annotation results

图5中的红点为视觉图像目标特征点,也是切

割路径的重要组成部分,可以结合特征点设计激光切割机器人的切割路径,分析图5(a)可知,所提方法能够有效提取视觉图像目标特征点。分析图5(b)可知,在提取到视觉图像目标特征点之后,通过改进YOLOv5s实现激光切割机器人视觉图像目标粗定位,能够有效保证视觉图像目标粗定位质量。分析图5(c)可知,在视觉图像目标粗定位之后对于切割机器人标注目标特征进行对齐处理,能够保证特征对齐质量。分析图5(d)可知,在目标特征对齐后确定激光切割机器人视觉图像关键凸显区域帧并确定目标特征点,利用这些点能够组成精准的切割路径,保证切割的精准度。

为了进一步验证所提方法的有效性,将文献[4]方法、文献[5]方法、文献[6]方法作为实验对比方法,通过比较交并比、定位误差、标注速度以及标注效果四个指标完成对比实验。

图6为图像目标标注交并比实验结果,对比每10次实验结果均值。

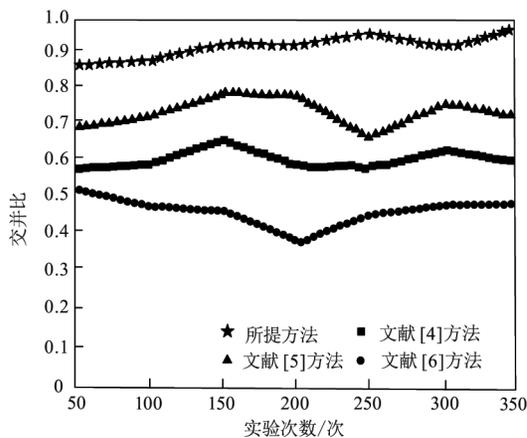


图6 图像目标标注交并比实验对比

Fig. 6 Comparison of image object annotation intersection and intersection experiment

从图6可以看出,文献[4]方法的图像目标标注交并比在0.58~0.65之间,文献[5]方法的图像目标标注交并比在0.67~0.79之间,文献[6]方法的图像目标标注交并比在0.36~0.50之间,所提方法的交并比数值始终高于0.85,和其他三个文献方法相比也处于领先地位,证明该方法可显著优化激光视觉图像标注性能,标注结果权威性与可靠性更强。

对比了四种方法的激光切割机器人视觉图像目标定位误差,结果如图7所示。

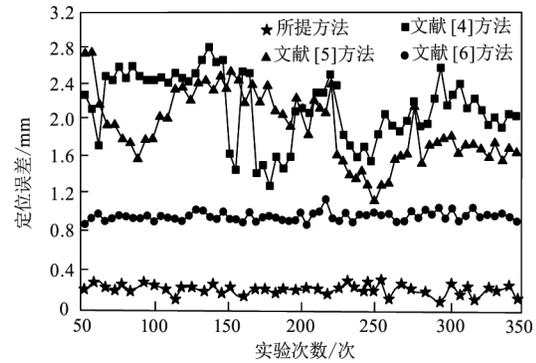


图7 目标定位误差实验对比

Fig. 7 Comparison of target positioning error experiments

分析图7中的数据可知,文献[4]方法的激光切割机器人视觉图像目标定位误差在1.31 mm~2.82 mm区间内变化,文献[5]方法的激光切割机器人视觉图像目标定位误差在1.17 mm~2.72 mm区间内变化,文献[6]方法的激光切割机器人视觉图像目标定位误差在0.84 mm~1.18 mm区间内变化,所提方法的激光切割机器人视觉图像目标定位误差在0.37 mm以下。与实验对比方法相比,所提方法在进行激光切割机器人视觉图像目标标注过程中,对于目标定位的误差最小,说明该方法具备较高的定位精度。

图像目标标注效率实验对比如表1所示。

表1 图像目标标注效率实验对比(单位:s/幅)

Tab. 1 Experimental comparison of image object annotation efficiency (Unit:s/image)

实验次数/次	文献[4]方法	文献[5]方法	文献[6]方法	所提方法
50	4.66	7.96	5.69	1.36
100	4.98	6.48	6.47	1.47
150	5.58	7.56	5.88	1.58
200	4.47	6.48	4.41	1.69
250	5.35	7.12	5.32	1.87
300	4.55	7.63	5.14	1.66
350	6.36	7.19	5.66	1.82

由表1可以看到,当实验次数为100次的情况下,文献[4]方法的图像目标标注效率为4.98s/幅,文献[5]方法的图像目标标注效率为6.48s/幅,文献[6]方法的图像目标标注效率为6.47s/幅,所提方法的图像目标标注效率为1.47s/幅;当实验次数为300次的情况下,文献[4]方法的图像目标标注效率为6.36s/幅,文献[5]方法的图像目标标注效率为7.19s/幅,文献[6]方法的图像目标标注效率

为 5.66 s/幅,所提方法的图像目标标注效率为 1.82 s/幅。与三个文献方法相比,所提方法的标注速度依旧是最快的,以此证明该方法在规定时间内可完成更多的图像标注任务。

以上几种方法的图像目标标注效果如图 8 所示。

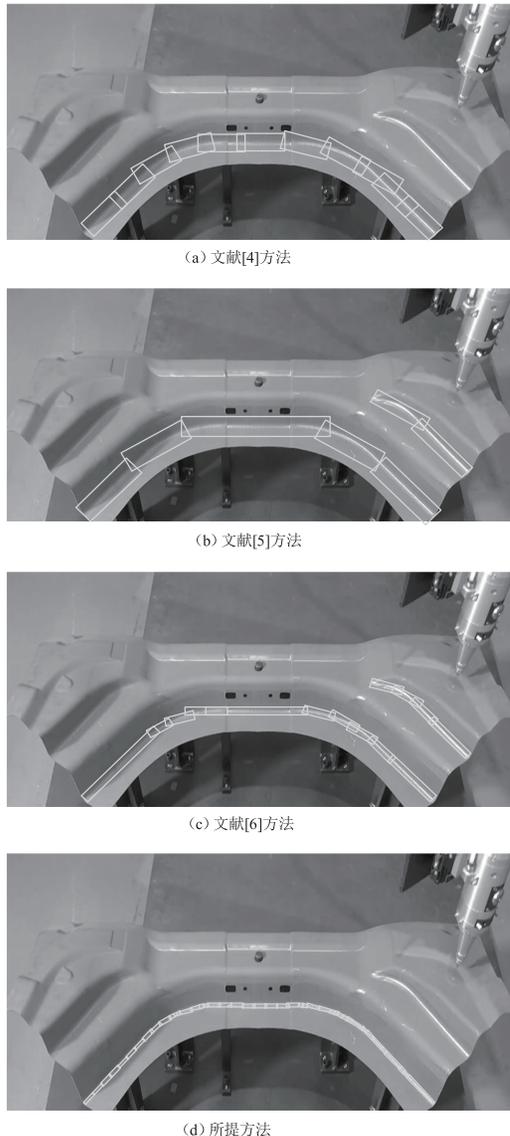


图 8 图像目标标注效果

Fig. 8 Image target annotation effect

分析图 8(a)可知,文献[4]方法在对于图像目标检测过程中,其图像目标标注范围较大,标注结果中非切割路径的范围较大,说明该方法的标注效果比较粗糙,实际应用效果并不好。分析图 8(b)可知,文献[5]方法在对于图像目标检测过程中,标注结果中非切割范围较文献[4]方法较大,且还将左上的非切割路径标注为切割路径,说明该方法存在错误标注的问题。分析图 8(c)可知,文献[6]方法

标注出来的非切割范围较文献[4]方法、文献[5]方法小,但是存在错误标注的问题,说明该方法的标注效果不高。分析图 8(d)可知,所提方法的图像目标标注结果中的非切割范围非常小,标注更为细致,且不存在错误标注的问题,说明该方法的激光切割机器人视觉图像目标标注效果更好。

## 5 结论

面向激光加工问题提出一种基于改进 YOLOv5s 的激光切割机器人视觉图像目标标注方法,通过改进 YOLOv5s 模型实现图像目标粗定位,使用目标特征对齐增强目标标注时效性,运用关键帧选择与半监督对抗学习完成图像标注任务。通过大量实验可以看出,所提方法可显著改善图像标注准确性,能够为激光切割机器人的自动化加工提供可行思路。

## 参考文献:

- [1] Qin Hao, Chen Jiyang, Lian Zhijie, et al. Robotic welding planning of medium and thick plates based on laser vision [J]. *Modular Machine Tool & Automatic Manufacturing Technique*, 2022(2): 128 - 132. (in Chinese)  
秦浩, 陈纪阳, 连智杰, 等. 基于激光视觉的机器人中厚板焊接规划[J]. *组合机床与自动化加工技术*, 2022(2): 128 - 132.
- [2] Chen Xinyu, Zhang Qingxin, Zhu Linlin, et al. The method of real time seam tracking for robotic welding system based on laser vision sensor [J]. *Laser & Infrared*, 2021, 51(4): 421 - 427. (in Chinese)  
陈新禹, 张庆新, 朱琳琳, 等. 基于激光视觉传感器的机器人实时焊缝跟踪方法[J]. *激光与红外*, 2021, 51(4): 421 - 427.
- [3] Zhang Yuning, Guan Junnan, Zhu Lina, et al. Automatic extraction method of welding trajectory from point cloud for robotic laser welding [J]. *Machine Tool & Hydraulics*, 2023, 51(11): 7 - 12. (in Chinese)  
张宇宁, 官俊楠, 朱丽娜, 等. 基于点云的机器人激光焊接轨迹自动提取方法[J]. *机床与液压*, 2023, 51(11): 7 - 12.
- [4] Wang Wei, Wang Li, Chen Chuixin. Target recognition technology of robot vision spectral image based on CNN [J]. *Laser Journal*, 2022, 43(3): 118 - 122. (in Chinese)  
王薇, 汪理, 陈吹信. 基于 CNN 的机器人视觉光谱图像目标识别技术[J]. *激光杂志*, 2022, 43(3): 118 - 122.

- [5] Huang Yimeng, Yi Yang. Robot object detection and localization based on deep learning[J]. Computer Engineering and Applications, 2020, 56(24): 181 - 187. (in Chinese)  
黄怡蒙, 易阳. 融合深度学习的机器人目标检测与定位[J]. 计算机工程与应用, 2020, 56(24): 181 - 187.
- [6] Guo Yuqing, Zeng Qingjun, Xia Nan, et al. Research on target recognition of autonomous underwater vehicle based on image enhancement[J]. China Measurement & Test, 2021, 47(11): 47 - 52. (in Chinese)  
郭雨青, 曾庆军, 夏楠, 等. 图像增强水下自主机器人目标识别研究[J]. 中国测试, 2021, 47(11): 47 - 52.
- [7] Wu Jianhua, Bai Yongfu, Xie Yongsheng, et al. Research on object detection algorithm of inspection robot based on multi-source information fusion[J]. Radio Engineering, 2023, 53(8): 1925 - 1933. (in Chinese)  
吴建华, 白永福, 谢永生, 等. 多源信息融合的巡检机器人目标检测算法研究[J]. 无线电工程, 2023, 53(8): 1925 - 1933.
- [8] Li Xiaopei, Zhang Yinbao, Li Yanpei, et al. An improved method of infrared image target detection based on YOLOv5s[J]. Laser & Infrared, 2023, 53(7): 1043 - 1051. (in Chinese)  
李晓佩, 张寅宝, 李严培, 等. 一种基于 YOLOv5s 的红外图像目标检测改进算法[J]. 激光与红外, 2023, 53(7): 1043 - 1051.
- [9] Chen Xu, Peng Dongliang, Gu Yu. Real-time object detection for UAV images based on improved YOLOv5s[J]. Opto-Electronic Engineering, 2022, 49(3): 69 - 81. (in Chinese)  
陈旭, 彭冬亮, 谷雨. 基于改进 YOLOv5s 的无人机图像实时目标检测[J]. 光电工程, 2022, 49(3): 69 - 81.