

基于 YOLOv5 视觉感知的实时叉车驾驶操作行为识别方法*

陈友升¹ 周介祺¹ 梁敏健^{2,3} 张大伟³ 刘桂雄¹

(1.华南理工大学机械与汽车工程学院, 广东 广州 510640 2.广东省特种设备检测研究院珠海检测院, 广东 珠海 519002 3.阿拉善盟特种设备检验所, 内蒙古 阿拉善盟 750306)

摘要: 驾驶员操作信息可反映驾驶情况和操作水平, 有助于驾驶安全检测。以叉车为例, 研究一种基于 YOLOv5 视觉感知的实时叉车驾驶操作行为识别方法。首先, 提出该方法的整体架构; 然后, 研究基于操作特征的实时叉车驾驶操作行为识别方法; 最后, 在一台合力 K30 柴油叉车上搭建试验平台进行测试。测试结果表明: 本方法操作特征识别率达 100%、操作行为识别正确率达 98%, 具有准确性好、实时识别的特点, 可拓展应用于其他车辆操作检测。

关键词: 深度学习; 目标检测; 行为识别; 驾驶操作

中图分类号: TH391.4

文献标识码: A

文章编号: 1674-2605(2021)03-0004-06

DOI: 10.3969/j.issn.1674-2605.2021.03.004

0 引言

驾驶员通过与操纵器交互获取驾驶操作信息, 以便监控驾驶状态^[1]、评估驾驶水平^[2]。目前, 获取驾驶操作信息的主要方法包括基于车辆信息感知、基于驾驶员生理信号反应和基于驾驶员操作行为感知等。其中基于车辆信息感知方法通过检测车辆行驶信息, 反推驾驶员操作行为, 如文献[3] (2015) 通过 MEMS 惯性传感器、汽车 OBD 系统, 完成驾驶员操作行为识别, 并分析驾驶操作与油耗之间的关系; 文献[4] (2018) 通过车辆 CAN 总线获取行驶信息, 经滤波、数据融合以及工况识别, 反推驾驶操作行为; 文献[5] (2020) 通过 DAS 系统采集行车信息, 使用一种基于双向长短记忆网络 (Bi-LSTM) 及全连接神经网络 (FC), 拓展神经网络检测模型, 借助行车数据在时间序列上发生突变的特征判断异常驾驶。基于驾驶员生理信号反应方法通过检测驾驶员脑电、心电、肌电、眼电、皮电等生理信号变化, 实现驾驶操作行为识别, 如文献[6] (2012) 介绍一种昏睡驾驶识别方法, 采用多传感器融合技术, 通过在驾驶员关键部位放置电极, 检测生物电情况判断驾驶员操作状态。基于驾驶员操作行为感知方法一般采用视觉传感器直接获取驾驶员操作行为进行判断识别, 如文献[7] (2018) 通过单

目视觉获取运动中驾驶员的视频流, 将驾驶员 8 个骨骼节点的三维坐标和预先建立的三维人体模型的二维投影进行匹配, 检测和识别驾驶员上半身正常、单手接听电话、疲劳、醉酒驾驶 4 种状态; 文献[8] (2018) 使用视觉传感器, 采用 LetNet, AlexNet 卷积神经网络分类网络, 完成驾驶员抽烟、玩手机等违规行为的有效识别; 文献[9] (2019) 研究一种基于 CNN 的非安全驾驶行为视觉传感识别方法, 平均识别率达 97.13%。综上所述: 基于车辆信息感知方法、基于驾驶员生理信号反应方法采用间接技术, 难以获取驾驶员真实操作的完整信息; 基于驾驶员操作行为感知方法, 采用视觉传感器, 可实时获取驾驶操作信息, 是目前驾驶操作行为识别的主流方法。

本文采用驾驶员操作行为感知方法, 结合深度学习的图像目标检测算法, 提出一种基于 YOLOv5 视觉感知的实时叉车驾驶操作行为识别方法, 利用工业相机实时识别叉车驾驶操作行为。

1 基于 YOLOv5 视觉感知的实时叉车驾驶操作行为识别方法

基于 YOLOv5 视觉感知的实时叉车驾驶操作行为识别方法包括: 1) 基于 YOLOv5 的实时叉车驾驶操作特征提取方法; 2) 基于驾驶操作逻辑的实时叉

* 基金项目: 国家市场监督管理总局技术保障专项项目 (2019YJ014); 广东省特种设备检测研究院科技项目 (2020JD10)。

车驾驶操作行为识别方法，框图如图 1 所示。通过视觉传感器结合深度学习目标检测算法，获取驾驶操作

特征，建立手部与各个操纵器之间的控制模型，实现驾驶操作行为识别。

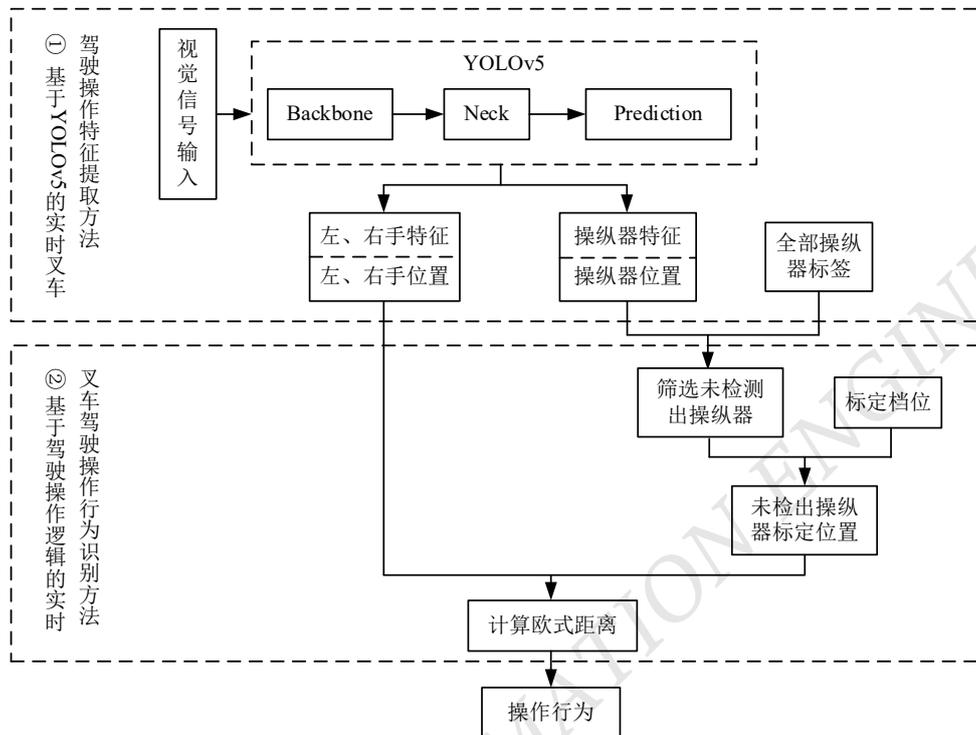


图 1 基于 YOLOv5 视觉感知的实时叉车驾驶操作行为识别方法框图

1.1 基于 YOLOv5 的实时叉车驾驶操作特征提取方法

常见的深度学习目标检测算法，如 Faster R-CNN 将检测分为候选框提取、分类 2 个过程，在模型训练过程中需对 RPN, fast R-CNN 进行反复训练，计算量

大、检测时间长。2016 年 Joseph Redmon 等提出基于回归 YOLO 目标检测算法^[10]。该算法结构简单、检测速度快。YOLO 模型经过不断改进与性能提高，2019 年在开源社区提出 YOLOv5 模型。YOLOv5 网络结构如图 2 所示。

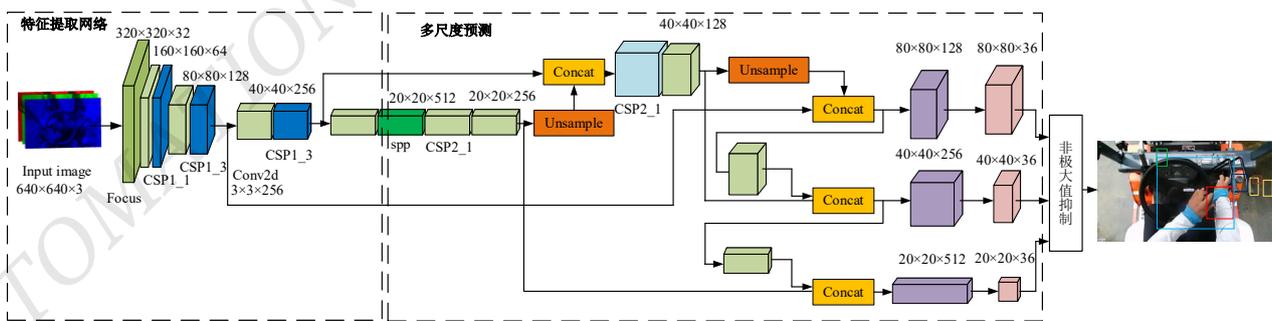


图 2 YOLOv5 网络结构

首先，对输入的叉车驾驶图片进行特征提取，通过 Focus 及 CSP 结构压缩特征图，通过 FPN 网络进行特征信息的传递融合，从而得到预测后的特征图；

然后，通过 PAN 结构提升预测性能；最后，在预训练好的 YOLOv5 网络中，对叉车关键部件及人体操作关键位置进行多尺度预测，输出包括双手、操作档

杆、方向盘等多个目标的定位框。

YOLOv5 模型的损失 L_{YOLOv5} 由回归损失 L_{box} 、正样本分类损失 L_{cls} 和正负样本的前景背景预测分类损失 L_{obj} 组成：

$$L_{YOLOv5} = L_{box} + L_{cls} + L_{obj} \quad (1)$$

将人工标注数据集输入到 YOLOv5 模型，通过迭代减少 L_{YOLOv5} ，使其收敛到足够小，模型训练完成。相机采集实时驾驶行为视频输入到已训练模型中，即可完成实时目标检测。

1.2 基于驾驶操作逻辑的实时叉车驾驶操作行为识别方法

驾驶员通过手部操作控制叉车作业，检测叉车驾驶操作行为需建立手部与各个操纵器之间的控制模型，左右手控制对象及操作形式如表 1 所示。

表 1 左右手控制对象及操作形式

名称	控制对象	标签名	档位状态	操作形式	自动保持
左手 (LeftHand)	方向盘	SteerWheel	转角	旋转	是
	驻车制动	Brake	松制动 合制动	移动	是
右手 (RightHand)	方向盘	SteerWheel	转角	旋转	是
	叉架升降操纵器	ForkHeight	空挡 上升 下降	移动	否
	叉架倾斜操纵器	ForkLean	空挡 前倾 后仰	移动	否
	叉车快慢操纵器	ForkliftSpeed	空挡 快速 慢速	移动	是
	叉车方向操纵器	ForkliftDirection	空挡 前进 后退	移动	是

在规范驾驶过程中，左手负责控制方向盘与驻车制动；右手负责控制方向盘、叉架升降、叉架倾斜、叉车快慢、叉车方向操纵器。这些操纵器按照操作形式又分为移动式、旋转式 2 种，其操作逻辑不同，操作行为识别方法也不同。

驾驶员握住移动式操纵器手柄，用手掌推动其到达固定位置，完成档位切换操作。移动式操纵器操作

行为识别原理图如图 3 所示。

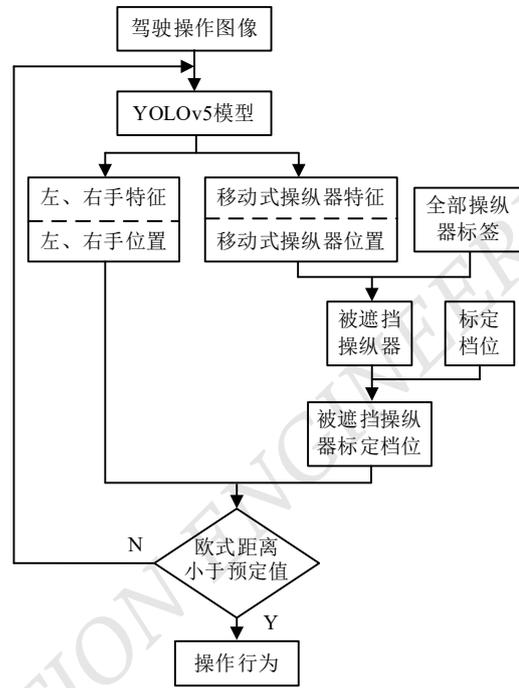


图 3 移动式操纵器操作行为识别原理图

在非驾驶状态对操纵器进行换挡操作，记录不同操纵器处于不同档位的图像，使用 YOLOv5 对图像进行目标检测，并提取图像特征。设图像像素坐标为 (u, v) ，特征图边界框为 $R_{bd}(u_{min}, v_{min}, u_{max}, v_{max})$ ，形心坐标为 $P_{pos} \left(\frac{u_{min} + v_{min}}{2}, \frac{u_{max} + v_{max}}{2} \right)$ ，可得到操纵器 n 在不同档位 m 下的标定档位集合 H_{pos-n} ，建立操纵器档位集合 Φ ：

$$H_{pos-n} : \{P_{pos-n}^1, \dots, P_{pos-n}^m, \dots, P_{pos-n}^M\} \quad (2)$$

$$\Phi : \{H_{pos-1}, \dots, H_{pos-n}, \dots, H_{pos-N}\}$$

可获得全部操纵器标签集合 Ω_0 。

采集驾驶过程中的操作行为视频，逐帧输入到已训练的 YOLOv5 模型；识别并提取当前图像中的操特征，可得到当前时刻图像中所有控制器标签集合 Ω ；通过补集操作，得到被遮挡档位集合 $\Delta\Omega$ ；根据被遮挡标签可得被遮挡操纵器对应的标定档位集合 $\Delta\Phi$ 。

提取当前图像中的手部特征、边界框 $R_{LeftHand}$ ， $R_{RightHand}$ ，得到其形心坐标 $P_{LeftHand}$ ， $P_{RightHand}$ ，计算手部 (X_{Hand}) 形心与标定档位间的欧式距离 d_m ：

$$d_m = \sqrt{(u_{X_{\text{Hand}}} - u_{\text{pos-}n}^m)^2 + (v_{X_{\text{Hand}}} - v_{\text{pos-}n}^m)^2} \quad (3)$$

针对不同的操纵器采用不同的安全距离 L_i , 当 d_m 小于安全距离 L_i 时, 即认为在进行操纵器 n 切换到 m 档位的操作。叉车升降操纵器操作识别过程图如图 4 所示。

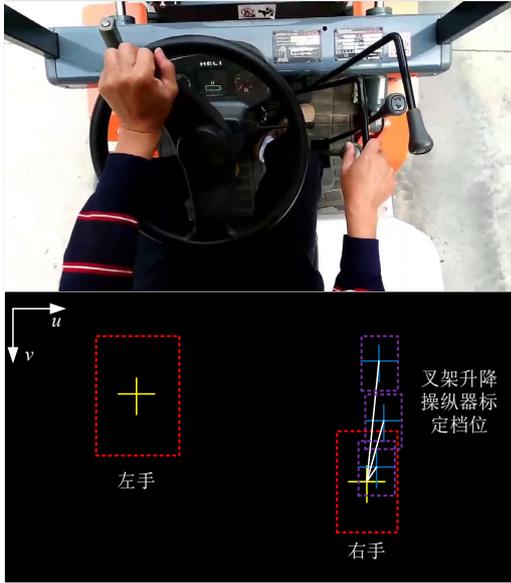


图 4 叉车升降操纵器操作识别过程图

方向盘是叉车上唯一的旋转式操纵器, 在规范操作中, 左手握住方向盘手柄旋转完成叉车转向动作, 识别方向盘操作行为即识别某一段时间内方向盘的旋转方向和转角。方向盘操作行为识别示意图如图 5 所示。设顺时针为正方向, 初始为 0 时刻, 通过 YOLOv5 模型可获得第 i 时刻左手形心 P_{LeftHand}^i 和方向盘形心 $P_{\text{steerwheel}}^i$ 、第 $i-1$ 时刻左手形心 $P_{\text{LeftHand}}^{i-1}$, 建立起向量 $V_i, V_{i-1} (V_i \neq V_{i-1})$, 令 LH, SW 为 LeftHand, SteerWheel 的简写, 引入转角方向系数 α :

$$\alpha = \begin{cases} 1, & \text{if } \begin{cases} (u_{\text{LH}}^i < u_{\text{LH}}^{i-1} \ \& \ v_{\text{LH}}^i, v_{\text{LH}}^{i-1} < v_{\text{SW}}^i) \\ \text{or } (u_{\text{LH}}^i, u_{\text{LH}}^{i-1} < u_{\text{SW}}^i \ \& \ v_{\text{LH}}^i > v_{\text{LH}}^{i-1}) \\ \text{or } (u_{\text{LH}}^i > u_{\text{LH}}^{i-1} \ \& \ v_{\text{LH}}^i, v_{\text{LH}}^{i-1} > v_{\text{SW}}^i) \\ \text{or } (u_{\text{LH}}^i, u_{\text{LH}}^{i-1} > u_{\text{SW}}^i \ \& \ v_{\text{LH}}^i < v_{\text{LH}}^{i-1}) \end{cases} \\ -1, & \text{else} \end{cases} \quad (4)$$

则在 $i-1$ 到 i ($i=1,2,3,\dots$) 时刻内, 转角 $\Delta\theta_i$ (由于时间间隔小, $\Delta\theta_i \in [0, \pi)$) 为

$$\Delta\theta_i = \alpha \arccos \frac{V_i V_{i-1}}{|V_i| \cdot |V_{i-1}|} \quad (5)$$

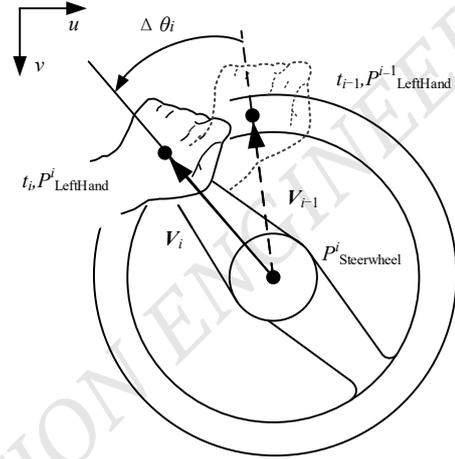


图 5 方向盘操作行为识别示意图

2 试验验证

在广东省某叉车培训场地进行试验。试验设备为合力 K30 柴油叉车, 视频采集设备为某品牌 1080 P 相机搭配 6 mm 广角镜头, 图像处理终端为搭载有 GTX 1080Ti GPU 的笔记本电脑。试验设备布置示意图如图 6 所示, 相机光轴垂直于方向盘平面。试验时, 采集 8 段共 30 min 的叉车操作视频, 使用聚类算法提取操作行为图片 500 张, 并进行标注构建模型训练数据集, 将训练集输入到模型中进行模型训练。

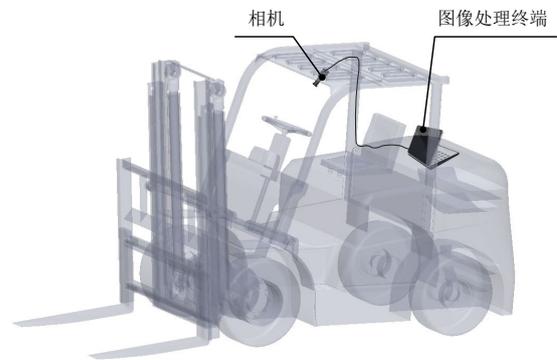


图 6 试验设备布置示意图

在叉车启动前,采集档位标定图片 30 张,并进行各操纵器档位标定。驾驶员驾驶叉车进行实操考核,相机实时采集驾驶操作行为图像,并输入到已训练的模型进行特征检测与操作行为识别。本实验共进行 3

段实操考核,实时检测视频共 10 min,处理图像约 18000 帧,依据每帧操作识别实验结果统计得出,操作特征检出率达 100%,操作行为识别率达 98%。叉车驾驶操作行为识别效果图如图 7 所示。



图 7 叉车驾驶操作识别效果图

3 结论

1) 基于 YOLOv5 的实时叉车驾驶操作特征提取方法,使用工业相机获取叉车驾驶实时操作行为图像,通过 YOLOv5 深度学习模型获取操纵器和驾驶员特征,识别定位叉车关键部件及驾驶员操作关键点坐标建立叉车驾驶员操作行为模型;

2) 基于驾驶操作逻辑的实时叉车驾驶操作行为识别方法,将图像特征信息输入驾驶操作逻辑算法,得到驾驶员当前操作信息,操作特征识别率达 100%,操作行为识别正确率达 98%。

下一步考虑对操作特征提取和操作行为识别算法进行轻量化处理,设计人工智能的嵌入式系统,将叉车驾驶行为检测系统产品化,实现叉车驾驶操作的自动化检测。

参考文献

- [1] 李聪.车辆驾驶状态监控预警系统设计[J].信息与电脑(理论版),2016(15):131-133,142.
- [2] 崔悦,石旭东,于之靖.机场特种车辆驾驶人员操作水平评价

系统[C]//第七届全国交通运输领域青年学术会议,2007.

- [3] 封宇华,杨拥民,杜凯,等.汽车驾驶行为与行驶状态的数据监测系统设计[J].单片机与嵌入式系统应用,2015,15(12): 54-57.
- [4] 杨新桦,刘宝健,张亮,等.一种驾驶员驾驶水平评价系统及其开发方法[J].重庆理工大学学报(自然科学),2018,32(9):1-5.
- [5] 惠飞,郭静,贾硕,等.基于双向长短记忆网络的异常驾驶行为检测[J].计算机工程与应用,2020,56(24):116-122.
- [6] SAHAYADHAS A, SUNDARAJ K, MURUGAPPAN M. Detecting driver drowsiness based on sensors: a review [J].Sensors, 2012,12(12):16937-16953.
- [7] 吴晨谋,方志军,黄正能.基于单目摄像头的主动式驾驶行为分析算法[J].山东大学学报(工学版),2018,48(5):69-76.
- [8] 李俊俊,杨华民,张澍裕,等.基于神经网络融合的司机违规行为识别[J].计算机应用与软件,2018,35(12):222-227,319.
- [9] 徐丹,代勇,纪军红.基于卷积神经网络的驾驶人行为识别方法研究[J].中国安全科学学报,2019,29(10):12-17.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[J]. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 779-788.

Real Time Forklift Driving Operation Behavior Recognition Method Based on YOLOv5 Visual Perception

Chen Yousheng¹ Zhou Jieqi¹ Liang Minjian^{2,3} Zhang Dawei³ Liu Guixiong¹

(1.School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China 2.Guangdong Institute of Special Equipment Inspection and Research Zhuhai Branch, Zhuhai 519002, China 3. Alxa Special Equipment Inspection Institute, Alxa 750306, China)

Abstract: Driver operation information can reflect the driving situation and operation level, which is helpful to driving safety detection. Taking forklift as an example, a real-time forklift driving behavior recognition method based on YOLOv5 visual perception is studied. Firstly, the overall framework of the method is proposed; Then, the real-time forklift driving behavior recognition method based on operation characteristics is studied; Finally, a test platform is built on a Heli K30 diesel forklift for testing. The test results show that the recognition rate of operation features is 100% and the correct rate of operation behavior recognition is 98%, which has the characteristics of good accuracy and real-time recognition, and can be extended to other vehicle operation detection.

Key words: deep learning; objection detection; behavior recognition; driving operation

作者简介:

陈友升, 男, 1996 年生, 硕士研究生, 主要研究方向: 先进传感与仪器。E-mail: 1445140@qq.com

刘桂雄 (通信作者), 男, 1968 年生, 教授, 博导, 主要研究方向: 先进传感与仪器。E-mail: megxliu@scut.edu.cn

~~~~~  
(上接第 14 页)

**Abstract:** In view of the problem that the vision system of the automatic fruit picker in the natural environment of the orchard is affected by the light, occlusion and other factors, which leads to the low detection accuracy of mature bayberry, An automatic detection method of mature bayberry based on machine vision is proposed. First of all, RGB camera is used to collect bayberry image in orchard environment. The input image is processed by contrast limited adaptive histogram equalization (CLAHE) based on Y component of YUV color space to compensate the influence of light on image quality. Then, the "G-R" and "2R-G-B" color difference methods are used to extract the background region and foreground region respectively. Based on the operations of morphology and region growth, the set of control points of foreground and background regions is established, and the region of mature bayberry fruit in natural environment is segmented automatically by OneCut segmentation algorithm. Finally, the watershed transform method and convex hull theory are used to segment the overlapped fruit, so that each arbutus fruit area can be located accurately. The experiment results show that the method can accurately detect the mature bayberry fruit in orchard environment, with an average precision of 93.5% and an average recall rate of 90.6%, which has a certain practical value.

**Key words:** orchard natural environment; mature bayberry detection; CLAHE; OneCut; watershed transformation

## 作者简介:

雷欢, 男, 1987 年生, 硕士, 工程师, 主要研究方向: 机器视觉、人工智能。

吴亮生, 男, 1987 年生, 在读硕士, 工程师, 主要研究方向: 机器视觉、人工智能。

焦泽昱, 男, 1991 年生, 博士, 助理研究员, 主要研究方向: 机器视觉、人工智能。

陈再励, 男, 1989 年生, 博士, 工程师, 主要研究方向: 安全工程、人工智能

马敬奇 (通信作者), 男, 1988 年生, 硕士, 工程师, 主要研究方向: 机器视觉、人工智能。E-mail: majingqiauto123@126.com

钟震宇, 男, 1971 年生, 博士, 研究员, 主要研究方向: 运动控制与机器视觉。