

本文引用格式: 赵彦龙,赵三伟,闫伟才,等.结合深度学习和边缘计算的作业人员专注性检测方法[J].自动化与信息工程,2023,44(5):46-51.

ZHAO Yanlong, ZHAO Sanwei, YAN Weicai, et al. Detection method of operator concentration combined with deep learning and edge computing[J]. Automation & Information Engineering, 2023,44(5):46-51.

结合深度学习和边缘计算的作业人员专注性检测方法

赵彦龙¹ 赵三伟² 闫伟才³ 钟震宇⁴

(1.内蒙古军区数据信息室, 内蒙古 呼和浩特 010051

2.武汉滨湖电子有限责任公司, 湖北 武汉 430205

3.中通服咨询设计研究院有限公司, 江苏 南京 210023

4.广东省科学院智能制造研究所/广东省现代控制技术重点实验室, 广东 广州 510070)

摘要: 在生产作业过程中, 作业人员注意力不集中是生产事故发生的主要诱因。针对现有的专注性研究方法过于依赖高性能计算设备, 存在部署困难和隐私数据泄露等问题, 提出一种结合深度学习和边缘计算的作业人员专注性检测方法。首先, 利用 YOLOv5 算法对使用手机、抽烟和喝水 3 种常见的分心行为进行实时检测; 然后, 采用轻量化人脸关键点检测网络和 Perclos 算法对作业人员的疲劳程度进行评估, 并对疲劳状态进行预警; 最后, 采用基于通道剪枝的压缩算法将分心行为检测网络和疲劳行为检测网络部署于低功耗的便携式边缘计算设备, 避免作业人员隐私和商业数据泄露的风险。实验结果表明, 该方法对疲劳行为和分心行为的检测准确率分别达 98.6% 和 99.8%, 满足实际的部署需求。

关键词: 专注性检测; YOLOv5 算法; 人脸关键点检测网络; Perclos 算法; 通道剪枝的压缩算法

中图分类号: TP391.4

文献标志码: A

文章编号: 1674-2605(2023)05-0007-06

DOI: 10.3969/j.issn.1674-2605.2023.05.007

Detection Method of Operator Concentration Combined with Deep Learning and Edge Computing

ZHAO Yanlong¹ ZHAO Sanwei² YAN Weicai³ ZHONG Zhenyu⁴

(1.Data Information Office, Inner Mongolia Military Region, Hohhot 010051, China

2.Wuhan Binhu Electronics Co., Ltd., Wuhan 430205, China

3.China Comservice Consulting Design Research Institute Co., Ltd., Nanjing 210023, China

4.Institute of Intelligent Manufacturing, Guangdong Academy of Sciences/Guangdong Key Laboratory of Modern Control Technology, Guangzhou 510070, China)

Abstract: During the production operation process, the lack of concentration of operators is the main cause of production accidents. In view of the existing focus research methods that rely too much on high-performance computing devices, which have problems such as deployment difficulties and privacy data leakage, a focus detection method for operators combining deep learning and edge computing is proposed. Firstly, the YOLOv5 algorithm is used to detect three common distracting behaviors: using a mobile phone, smoking, and drinking water in real-time; Then, a lightweight facial keypoint detection network and Perclos algorithm are used to evaluate the fatigue level of operators and provide early warning of fatigue status; Finally, the channel pruning based compression algorithm is used to deploy the distraction behavior detection network and fatigue behavior detection network to low-power portable

edge computing devices to avoid the risk of operators' privacy and commercial data leakage. The experimental results show that the detection accuracy of this method for fatigue behavior and distraction behavior reaches 98.6% and 99.8%, respectively, meeting the actual deployment requirements.

Keywords: concentration detection; YOLOv5 algorithm; face key point detection network; Perclos algorithm; compression algorithm for channel pruning

0 引言

随着生产作业日益复杂和自动化程度的不断提高,人机协作、人在回路等制造模式逐渐演化成智能制造的新范式,其中作业人员是生产制造过程中最为重要的角色之一^[1-3]。然而,面对高效率和高产出的生产任务,作业人员会发生注意力分散和疲劳等问题,不仅影响生产效率和产品质量,甚至可能引发生产事故,威胁作业人员的人身安全并导致精密昂贵的生产设备损坏,给生产作业带来巨大的安全隐患和经济损失。因此,及时地确定作业人员的专注性是保证人员安全、提升生产质量、避免意外事故的重要举措。

近年来,研究人员基于不同视角和方法对作业人员的专注性检测进行了一系列地研究。孙剑等^[4]提出一种基于自然驾驶数据的分心驾驶行为识别方法,该方法结合态势感知、决策树和随机森林对作业人员分心行为特征进行排序,并使用长短时记忆网络完成分心驾驶的分类识别,准确率达 89%。尹智帅等^[5]采用姿态估计方法检测视频中人体的姿态,并建立分心驾驶行为检测模型,该模型在实际采集的驾驶图像上取得 85.42% 的识别率。周建亮等^[6]通过测量分析建筑工人的心率、皮肤温度、舒张压、收缩压等 4 项生理指标,建立支持向量机模型用于疲劳检测,能够准确进行建筑工人的疲劳评估和预警。ZHAO 等^[7]利用结构光深度相机的深度特征,定位驾驶人员的空间位置,并完成驾驶分心行为检测,实验结果表明,深度相机能够有效改善光照噪声问题,提高识别效果。

虽然以往研究人员在专注性检测领域已取得显著的进展,但大部分研究聚焦于交通驾驶领域,对生产制造领域的研究较少,且研究方法存在以下局限:1) 功能单一,以往的研究方法只关注分心行为的检测,忽视了作业人员的疲劳程度与专注性的关联;2)

隐私暴露,目前多数研究通过建立大模型来提升检测的精度或效果,忽视了算法的易部署性和隐私性。

针对当前研究面临的局限性,本文提出一种结合深度学习和边缘计算的作业人员专注性检测方法。该方法由分心行为检测网络和疲劳行为检测网络两部分组成。其中,分心行为检测网络利用 YOLOv5 算法^[8-10]对使用手机、抽烟和喝水 3 种常见的分心行为进行实时检测;疲劳行为检测网络借助人脸关键点检测算法和 Perclos 算法^[11],对作业人员的疲劳程度进行评估,并对疲劳状态进行预警。此外,为避免作业人员隐私和商业数据泄露,采用基于通道剪枝的压缩算法对分心行为检测网络和疲劳行为检测网络进行压缩,并部署于边缘计算设备,实现离线环境下作业人员的实时专注性检测。

1 检测方法

结合深度学习和边缘计算的作业人员专注性检测方法流程如图 1 所示。

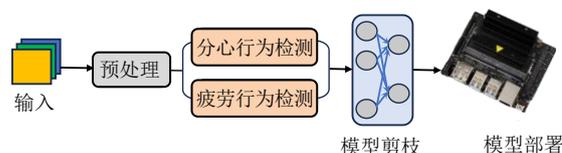


图 1 结合深度学习和边缘计算的作业人员专注性检测方法流程图

首先,将实时监测的作业人员视频帧通过亮度调整和批归一化操作进行预处理,减少图像噪声干扰;然后,通过分心行为检测网络和疲劳行为检测网络进行检测;最后,基于通道剪枝的压缩算法对分心行为检测网络和疲劳行为检测网络进行压缩,并部署于边缘计算设备,实现作业人员实时专注性检测。

1.1 分心行为检测

考虑到作业人员发生分心行为时,不可避免地导

致其专注性下降，影响正常的生产作业。因此，本文采用目标检测算法 YOLOv5 对工作场景下的作业人员进行分心行为实时检测。

1.1.1 YOLOv5

YOLOv5 是一种快速的 One-Stage 目标检测算法，由输入层、主干网络、颈部网络和输出层组成，网络结构如图 2 所示。

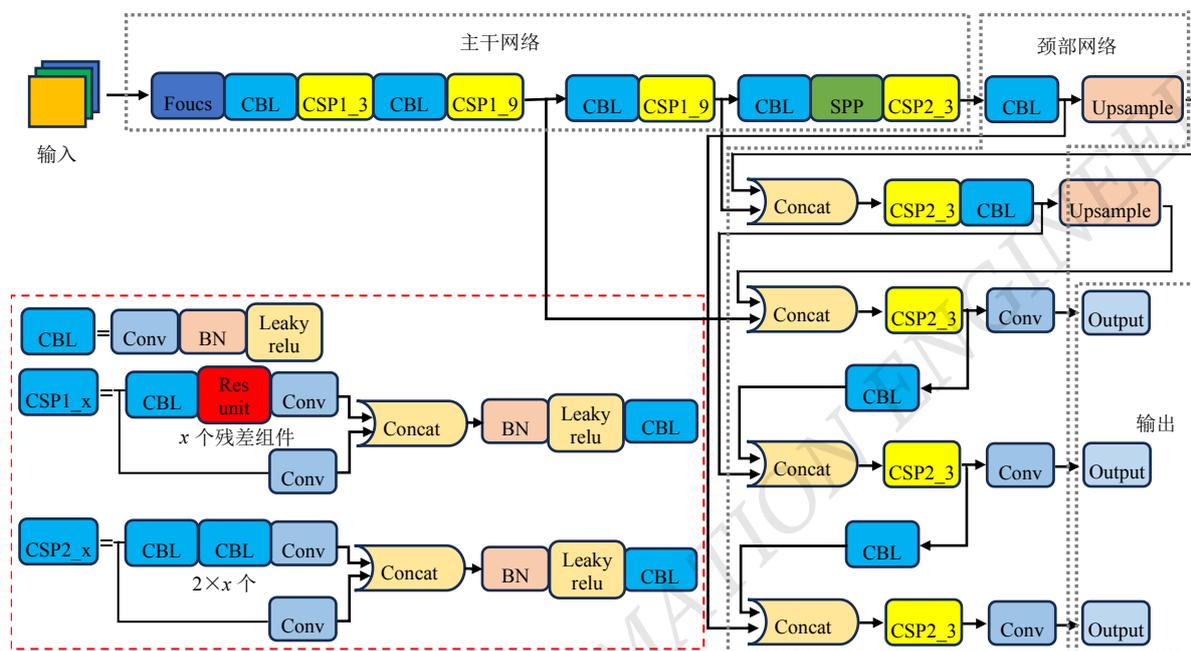


图 2 YOLOv5 算法结构

首先，输入待检测作业人员的图像；然后，主干网络对输入图像进行特征提取；接着，颈部网络对提取的特征进行多尺度特征融合；最后，输出层对融合后的特征进行分析并推理出检测结果。

考虑到生产作业空间的复杂性及不同工作场景下视频图像的尺寸不同，直接对视频图像进行裁剪易丢失关键信息，因此在主干网络中采用空间金字塔池化（spatial pyramid pooling, SPP）模块。SPP 模块先基于 3 个大小不同的池化核将模块输入的特征图划分为不同尺寸的网格，并对每个网格进行最大池化处理；再将池化结果堆叠成特定尺寸的特征图，从而避免因裁剪而导致重要的特征丢失。

为加强模型融合特征的能力，YOLOv5 算法在主干网络和输出层插入颈部网络。颈部网络融合了特征金字塔网络和路径聚合网络的优势，通过上采样方法将高层特征图与低层特征图融合，提升不同尺寸目标的检测能力。上采样方法完成自顶向下的语义特征传

递后，网络中缺少目标的定位特征信息，因此在特征金字塔网络之后引入路径聚合网络，实现自底向上的定位特征传递，融合底层的位置信息，加强网络的特征融合。输出层利用分类器将颈部网络获取的特征图进行分类，使用回归器对预测框进行修正微调，最终输出检测结果。

1.1.2 分心行为检测网络

本文基于 YOLOv5 算法构建分心行为检测网络的过程如下：

- 1) 使用大量具有分心行为的视频帧对 YOLOv5 算法进行训练，使该算法学习到分心行为的图像特征，并将这些图像特征与对应的行为关联；
- 2) 在测试过程中，利用训练好的 YOLOv5 算法对输入的视频帧进行判断，实现分心行为的实时检测。

1.2 疲劳行为检测

为评估作业人员的疲劳程度并及时发出疲劳警告，首先，采用基于方向梯度直方图（histogram of

oriented gradient, HOG) 的轻量化人脸关键点检测作业人员图像中 68 个人脸关键点, 包括眉毛 10 个, 鼻子 9 个, 眼睛 12 个, 嘴巴 20 个, 下巴 17 个; 然后, 基于 Perclos 算法根据 68 个人脸关键点分析疲劳行为并计算疲劳程度; 最后, 基于上述计算结果设计疲劳预警机制反馈预警信号。

1.2.1 轻量化人脸关键点检测网络

考虑到生产制造领域的作业空间复杂和活动区域有限, 为保证算法易部署, 采用结合 HOG 特征提取算法和支持向量机的人脸关键点检测网络。该网络具有轻量化、高实时性、易部署等特点。人脸关键点检测过程如下:

- 1) 将输入的人脸图像分解成 N 个尺寸固定的区域块, 并计算梯度信息;
- 2) 根据梯度方向将每个区域块的梯度信息划分为 9 个区间, 并计算梯度直方图;
- 3) 将每个区间梯度直方图特征输入到支持向量机进行训练和测试, 得到 68 个人脸的关键点。

1.2.2 Perclos 算法

基于人脸关键点, 利用眼睛及嘴巴的开合程度判断是否发生眨眼或打哈欠等疲劳性表征活动, 并利用 Perclos 算法量化疲劳程度。

- 1) 眨眼检测, V_{ear} 表示眼睛开合程度, 人睁开时, V_{ear} 在某个值域范围内波动; 人眼闭合时, V_{ear} 迅速下降, 若 V_{ear} 低于某个阈值, 设定眼睛处于闭合状态; 当 V_{ear} 由某个值迅速下降至小于阈值, 又迅速上升至大于该阈值时, 则判断为一次眨眼。为了分析作业人员的疲劳程度, 需统计单位时间内眼睛开合程度 V_{ear} 的数值变化, V_{ear} 的计算公式为

$$V_{\text{ear}} = \text{Norm} \left(\frac{P_{42,y} + P_{41,y} - P_{38,y} - P_{39,y}}{P_{40,x} - P_{37,x}} \right) \quad (1)$$

式中: $P_{i,y}$ 为第 i 个人脸特征点的 y 坐标值, $P_{i,x}$ 为第 i 个人脸特征点的 x 坐标值, Norm 为最大最小归一化。

- 2) 打哈欠检测, V_{mar} 表示嘴巴活动值, 嘴巴开

合程度的原理同眼睛相同, 其计算公式为

$$V_{\text{mar}} = \text{Norm} \left(\frac{P_{59,y} + P_{57,y} - P_{51,y} - P_{53,y}}{P_{55,x} - P_{49,x}} \right) \quad (2)$$

基于每一个视频帧中眼睛和嘴巴的开合程度, 利用 Perclos 算法量化作业人员的疲劳程度, 计算公式为

$$\text{Perclos} = \frac{N_{\text{ear}} + N_{\text{mar}}}{T} \times 100\% \quad (3)$$

式中: T 为检测时间的总帧数, N_{ear} 和 N_{mar} 分别为检测时间段内眼睛和嘴巴的闭合程度超过阈值的帧数。

1.2.3 疲劳预警机制

当眼睛开合程度达到瞳孔面积的 80% 时, 认为完成一次闭眼; 当嘴巴开合程度达到 80% 时, 认为完成一次打哈欠。统计每一帧中闭眼和打哈欠的帧数, 若 Perclos 值超过 70%, 则认为已达到疲劳阈值, 发出疲劳警告; 否则, 持续输出 Perclos 的量化结果, 疲劳预警流程图如图 3 所示。

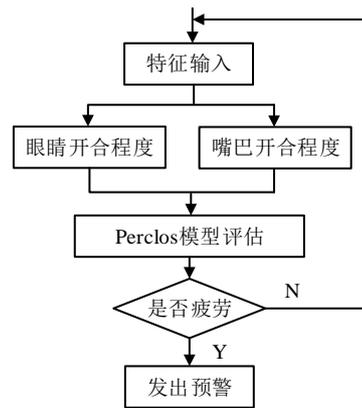


图 3 疲劳预警流程图

1.3 隐私保护策略

为了保护作业人员隐私和商业数据, 本文通过对分心行为检测网络和疲劳行为检测网络进行通道剪枝, 完成模型压缩, 以便部署于边缘计算设备, 实现安全、快速、稳定的专注性检测。

1.3.1 模型剪枝与压缩

在卷积神经网络中, 比例因子 γ 表示批归一化层

的权值, γ 越小, 表示该通道对最终结果的影响越小, 该通道的参数为冗余参数。因此, 本文采用通道剪枝的压缩算法对分心行为检测网络和疲劳行为检测网络进行压缩, 通道剪枝过程如图 4 所示。

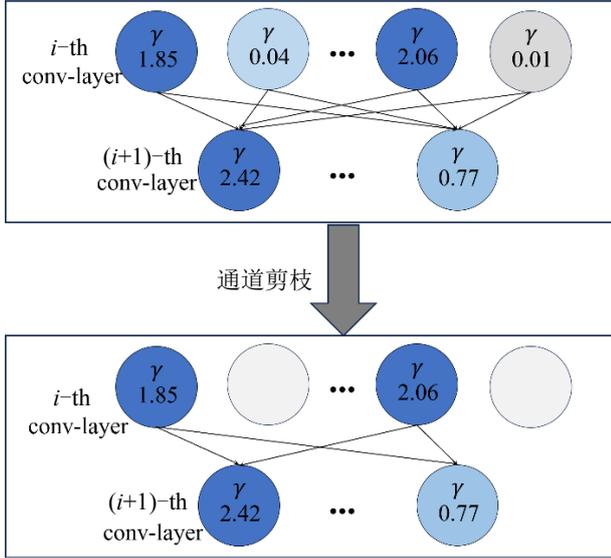


图 4 通道剪枝过程图

当分心行为检测网络和疲劳行为检测网络训练完成后, 通过稀疏训练对通道比例因子进行两极化。在通道的比例因子接近零时, 表明此通道对检测结果的影响可忽略, 则去除该通道。图 4 中每个圆圈表示卷积层的一个通道, 根据预先设定的剪枝率 p , 确定每个卷积层的剪枝阈值 θ , 剪枝过程为

$$\theta = \text{sort}_p(\Gamma) \quad (4)$$

式中: Γ 为每个卷积层中通道卷积因子的集合, $\text{sort}_p(\cdot)$ 为对通道比例因子按升序排序。

分心行为检测网络和疲劳行为检测网络经过剪枝后, 可以显著减少参数个数, 有效减少过拟合; 同时会出现精度损失, 可通过精度微调进行恢复。

1.3.2 边缘计算设备部署

在边缘侧部署设备时, 设备的体积、质量、成本是影响作业空间适用性的关键因素。本文采用 Nvidia Jetson Nano 为边缘计算设备部署。该设备搭载 128 个 Nvidia CUDA 核心, 四核 ARMA57 处理器, 2 GB 运

行内存, 功耗仅为 5~10 W, 上述特性满足部署需求。

2 实验结果和讨论

2.1 实验设置

分别采集抽烟、喝水和打电话 3 种分心行为的图像 726 幅、954 幅、865 幅, 并使用 Labelme 进行图像标注, 将每种分心行为图像的 80% 作为训练集, 20% 作为测试集。

分心行为检测网络和疲劳行为检测网络的训练和剪枝压缩的学习率均为 1×10^{-3} , 训练迭代次数均为 150, 采用随机梯度下降算法对网络进行参数优化。

在硬件配置方面, 采用 Nvidia DGX-1 图像深度学习训练平台, 该平台部署 8 张型号为 P100 的 GPU 处理器, 运行内存为 16 GB。所有实验均采用 Ubuntu18.04 操作系统, 基于 Pytorch 框架、OpenCV、NumPy 等开源库完成。

2.2 评价指标

为了评估本文方法的有效性, 分别采用准确率 (AP)、模型尺寸和推理速度 (frame per second, FPS) 作为评价指标, 其中准确率计算公式为

$$AP = \frac{TP}{TP + FP} \quad (5)$$

式中: TP 为正确识别分心行为或疲劳行为的视频帧数, FP 为将正常行为的视频帧误识别成分心行为或疲劳行为的帧数。

2.3 实验结果与讨论

为验证本文方法性能的优越性, 分别采用 Faster RCNN (F-RCNN) [12]、YOLOv3 [13]、YOLOv5l 三种方法与本文方法进行对比实验。在测试集上的实验结果如表 1 所示。

表 1 4 种方法的实验性能对比

方法	尺寸/MB	AP-疲劳	AP-分心	FPS/(f/s)
F-RCNN	108.0	97.1%	96.1%	13.2
YOLOv3	236.0	97.4%	98.2%	43.8
YOLOv5l	89.6	98.1%	98.9%	32.0
本文方法	8.3	98.6%	99.8%	40.5

由表 1 可知, 本文方法经过模型压缩后, 尺寸仅为 8.3 MB, 疲劳行为检测和分心行为检测的准确率分别为 98.6%、99.8%, 比 F-RCNN、YOLOv3、YOLOv5l 取得更优的性能。因此, 本文方法不仅在模型的大小、精度上满足实际应用需求, 而且部署在边缘计算设备的推理速度也满足工业生产要求。

3 结论

本文针对生产作业过程中, 作业人员注意力不集中导致生产事故的问题, 提出一种结合深度学习和边缘计算的作业人员专注性检测方法。该方法通过分心行为检测网络和疲劳行为检测网络, 实现作业人员的实时专注性评估, 并基于通道剪枝的压缩算法将模型部署于边缘计算设备上, 避免隐私泄露。

参考文献

- [1] 黄海丰, 刘培森, 李擎, 等. 协作机器人智能控制与人机交互研究综述[J]. 工程科学学报, 2022, 44(4): 780-791.
- [2] 游青山, 冉霞. 基于机器视觉的矿井作业人员行为监测及违章识别系统[J]. 自动化与信息工程, 2021, 42(4): 20-24.
- [3] 于云峰, 金葆青. 机械自动化制造系统的思考探析和构建[J]. 现代制造技术与装备, 2021, 57(9): 184-186.
- [4] 孙剑, 张一豪, 王俊骅. 基于自然驾驶数据的分心驾驶行为识别方法[J]. 中国公路学报, 2020, 33(9): 225-235.

- [5] 尹智帅, 钟恕, 聂琳真, 等. 基于人体姿态估计的分心驾驶行为检测[J]. 中国公路学报, 2022, 35(6): 312-323.
- [6] 周建亮, 陈玮, 范丽萍. 基于生理指标的建筑工人攀登作业疲劳实验研究[J]. 中国安全生产科学技术, 2023, 19(3): 195-202.
- [7] ZHAO M, BEURIER G, WANG H, et al. In vehicle driver postural monitoring using a depth camera kinect[R]. SAE Technical Paper, 2018.
- [8] Jocher G, Chaurasia A, Stoken A, et al. ultralytics/yolov5: v7. 0-yolov5 sota realtime instance segmentation[J]. Zenodo, 2022.
- [9] 张青春, 吴峥, 周玲, 等. 基于改进 YOLOv5 的车辆与行人目标识别方法[J]. 中国测试, 2023, 49(7): 82-88.
- [10] 郑楚伟, 林辉, 吴晓明, 等. 基于自适应空间特征融合的 YOLOv5 安全帽检测系统设计[J]. 机电工程技术, 2022, 51(9): 37-42.
- [11] SOMMER D, GOLZ M. Evaluation of PERCLOS based current fatigue monitoring technologies[C]//2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE, 2010: 4456-4459.
- [12] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.
- [13] LI C, WANG R, LI J, et al. Face detection based on YOLOv3[C]//Recent Trends in Intelligent Computing, Communication and Devices: Proceedings of ICCD 2018. Springer Singapore, 2020: 277-284.

作者简介:

赵彦龙, 男, 1984 年生, 硕士研究生, 工程师, 主要研究方向: 计算机网络、人工智能、大数据。E-mail: 757477184@qq.com

赵三伟, 男, 1984 年生, 硕士研究生, 高级工程师, 主要研究方向: 雷达信号处理、抗干扰技术、目标分类识别。

E-mail: 277314744@qq.com

闫伟才(通信作者), 男, 1985 年生, 硕士研究生, 高级工程师, 主要研究方向: 无线通信。E-mail: yanweicai@foxmail.com

钟震宇, 男, 1971 年生, 博士研究生, 研究员, 主要研究方向: 深度学习、人工智能、大数据。E-mail: zy.zhong@giim.ac.cn

(上接第 21 页)

作者简介:

王红军, 女, 1966 年生, 博士研究生, 教授, 主要研究方向: 农业机器人、智能设计。E-mail: xtwhj@scau.edu.cn

邹伟锐, 男, 1998 年生, 硕士研究生, 主要研究方向: 虚拟现实。E-mail: 904507581@qq.com

谢启旋, 男, 1997 年生, 硕士研究生, 工程师, 主要研究方向: 机器视觉。E-mail: 896428001@qq.com

郑文和, 男, 1985 年生, 本科, 工程师, 主要研究方向: 智能装备设计。E-mail: justy.zheng@byd.com