

## 基于 YOLOv5 的姿态交互球类陪练机器人\*

曾杨吉, 刘自红, 蔡 勇, 郭星辰, 莫金龙

(西南科技大学 制造科学与工程学院, 四川 绵阳 621000)

**摘 要:** 针对当下球类陪练机器人人机交互能力不足的问题, 提出一种基于树莓派和 YOLOv5 目标检测算法的新型人机交互模式, 使机器人实现前进、后退、左移、右移、抛球、踢球 6 种不同的动作; 通过对在 3 种不同环境(室内、室外晴天、室外阴天)下搜集的人体姿态数据集进行标定、训练后, 得到 6 种姿态在 3 种环境中测试集上的识别准确率分别为: 室内 96.33%、室外晴天 95%、室外阴天 94.3%。相比基于特征匹配和其他利用手势等小目标检测的算法, 基于该算法的机器人具有更高的检测速度和准确性, 使机器人更加智能化。

**关键词:** YOLOv5 算法; 姿态识别; 球类陪练机器人; 树莓派; STM32 单片机

中图分类号: TN92; TH39

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.211736

中文引用格式: 曾杨吉, 刘自红, 蔡勇, 等. 基于 YOLOv5 的姿态交互球类陪练机器人[J]. 电子技术应用, 2022, 48(1): 76-79.

英文引用格式: Zeng Yangji, Liu Zihong, Cai Yong, et al. An interactive ball training partner robot based on YOLOv5[J]. Application of Electronic Technique, 2022, 48(1): 76-79.

## An interactive ball training partner robot based on YOLOv5

Zeng Yangji, Liu Zihong, Cai Yong, Guo Xingchen, Mo Jinlong

(School of Manufacturing Science and Engineering, Southwest University of Science and Technology, Mianyang 621000, China)

**Abstract:** In order to solve the problem of insufficient human-computer interaction ability of ball training partner robots, a new human-computer interaction mode based on Raspberry Pi and YOLOv5 algorithm was proposed, which enabled the robot to realize six different actions: forward, backward, left, right, throwing the ball, and kicking the ball. After calibrating and training the data sets collected in three different environments(indoor, outdoor sunny day and outdoor cloudy day), the recognition accuracy of the six poses in the test set under three different environments is 96.33% indoor, 95% outdoor sunny day, and 94.3% outdoor cloudy day, respectively. Compared with other algorithms based on feature matching and small target detection using gestures, the robot has higher detection speed and accuracy, which makes the robot more intelligent.

**Key words:** YOLOv5 algorithm; posture recognition; ball training partner robot; Raspberry Pi; STM32 MCU

## 0 引言

球类陪练机器人以其机动性好、针对性强、训练数据可量化等特点已在多个体育领域有广泛的应用, 目前对陪练机器人的研究大多集中于机器人控制方法<sup>[1]</sup>和机械结构的设计、优化<sup>[2-3]</sup>, 以提高机器人的多功能性和可控性为主。但是对于训练内容更加复杂的运动(如篮球、足球、橄榄球等), 还需要机器人具备更高的人机交互能力以发挥出最优性能。

传统的人机交互一般依赖于遥控器、键盘鼠标等外部设备, 存在一定的固有缺陷, 而基于图像处理的目标识别方法可实现与机器人的实时交互, 进一步提高机器人智能化水平。在基于图像处理的姿态识别研究中, 主要集中于以下几个方面: (1) 传感器: 基于惯性传感器、FDC2214 电容传感器等<sup>[4-5]</sup>; (2) 机器学习: 基于支持向量

机(SVM)、长短期记忆神经网络(LSTM)等<sup>[6-7]</sup>; (3) 深度学习模型: 基于 YOLO、卷积神经网络(CNN)等的方法在识别问题中也得到了广泛应用<sup>[8-9]</sup>。YOLO 系列算法由 Redmon 等人在 2015 年提出<sup>[10]</sup>, 后续不断发展为 YOLOv2<sup>[11]</sup>、YOLOv3<sup>[12]</sup>、YOLOv4<sup>[13]</sup>和最新的 YOLOv5。该算法的核心思想在于将目标检测统一为回归问题进行求解, 相较于同一精确度下采用分类+回归方法的 Faster R-CNN 等机器学习方法, 检测速度提升了约 4 倍。YOLOv5 是 YOLO 系列最新的算法, 是目前检测领域综合性最好的算法之一。

综上所述, 为了进一步提高球类陪练机器人的智能化水平, 实现与运动员的实时交互, 本文提出了一种新型人机交互模式, 即基于树莓派的姿态识别方法, 以橄榄球陪练机器人为实验平台, 利用 YOLOv5 目标检测算法对运动员特定动作进行识别, 以识别结果控制机器人的全方位置移动以及各机构动作, 使机器人与运动员具备

\* 基金项目: 西南科技大学博士基金(20zx7148); 西南科技大学研究生创新基金(20ycx0066)

更好的默契,提高训练质量。

## 1 系统方案介绍

### 1.1 实验平台介绍

实验机器人作为一种橄榄球陪练机器人,其主要机构包括:底盘机构、后射机构、前射机构;机器人通过STM32控制,在不同指令下完成前进、后退、左移、右移、前射踢球、后射抛球等一系列动作,机器人模型及实物如图1所示。

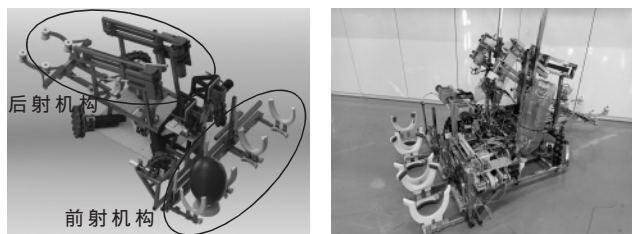


图1 机器人模型和实物图

### 1.2 控制方案设计

系统控制框架由摄像头、树莓派、STM32、电机及其驱动器和各传感器组成,如图2所示。摄像头作为图像采集模块,将运动员不同的姿态采集后发送至树莓派;树莓派将视频数据处理后通过串口通信将结果发送至STM32控制单元;STM32根据每种姿态识别结果计算对应的电机参数发送至电机驱动板完成电机驱动,同时机器人自身传感器(如激光、码盘、编码器等)采集机器人

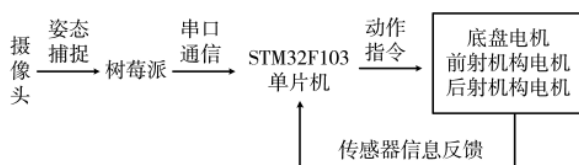


图2 控制框架图

自身状态,实时反馈至STM32形成闭环控制,实现对机器人基于姿态识别的精确控制。

## 2 YOLOv5 网络

YOLOv5 网络结构如图3所示,其继承了YOLOv4的网络架构,输入端采用mosaic数据增强、自适应锚框计算、自适应图片缩放。mosaic数据增强方法增强了对小目标的检测能力,有利于本实验中对于各种姿态中小特征的提取;Backbone组件采用Focus、CSP结构,其中Focus结构为YOLOv5较之前的版本的更新之处,其关键步骤在于将输入的图片进行切片操作;Neck组件采用FPN+PAN结构;输出端采用GIOU\_Loss方法,有效解决了边界框不重合问题<sup>[14]</sup>。YOLOv5与YOLOv4相比较:YOLOv5虽然在精确度上还不如YOLOv4,但是140帧/s的处理速度远远快于YOLOv4,对于一些需要在快速运动中完成目标识别的任务,YOLOv5的优势更加明显。

### 3 数据集制作

在本实验中,数据集由室内、室外晴天、室外阴天3种环境下采集的人体姿态图片组成,包括6种不同类型的姿态(向前、向后、向左、向右、前射、后射),每种姿态样本为300张,一组共1800张RGB图片,样本数据如图4所示。为了提高训练模型对不同环境下目标检测的

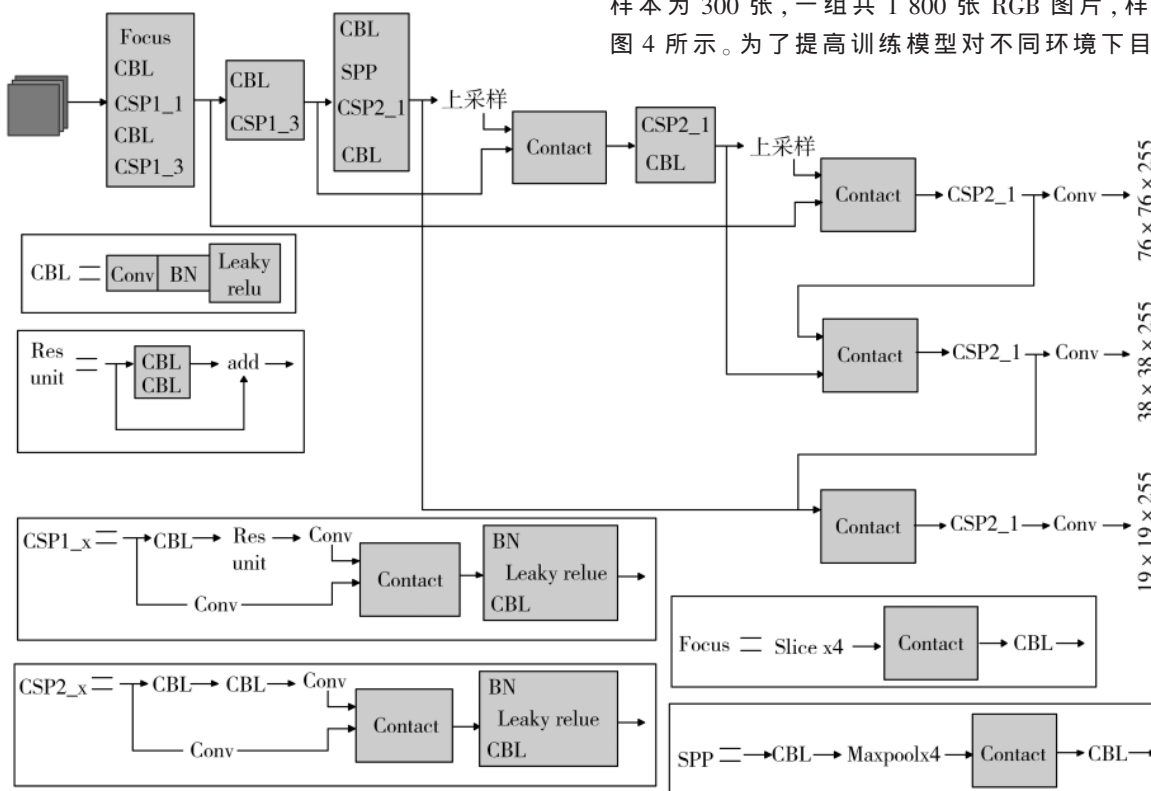
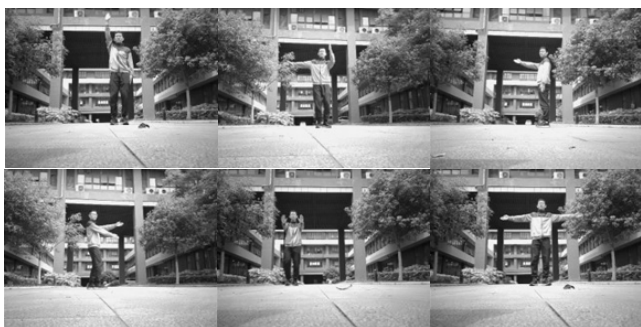


图3 YOLOv5 网络结构



(a)室内数据集



(b)室外晴天数据集



(c)室外阴天数据集

图4 姿态数据集样本

鲁棒性以及防止过拟合,通过缩放、旋转、对比度变化等数据增强方法扩充数据集生成新的样本,处理后,数据集中每种姿态图片为600张,一组共3600张,其中每种姿态训练集有480张,测试集有120张。在对数据集进行标定之前,将每一张图片首先进行归一化处理为 $416 \times 416$ 大小,随后采用labelImg标注软件手动框选图中人的姿态,最后导出为.txt文件。将处理好的数据集在Windows 10系统上选用PyTorch框架进行网络模型的训练,在训练阶段,batch设置为4,总迭代次数设置为300,模型训练后保存,在树莓派中调用此模型运行。

#### 4 实验验证与讨论

##### 4.1 视觉检测效果

测试集上的mAP在室内、室外晴天、室外阴天3种环境下分别可以达到96.33%、95%、94.3%,其中每种姿态在室内、室外晴天、室外阴天环境下的识别准确率分

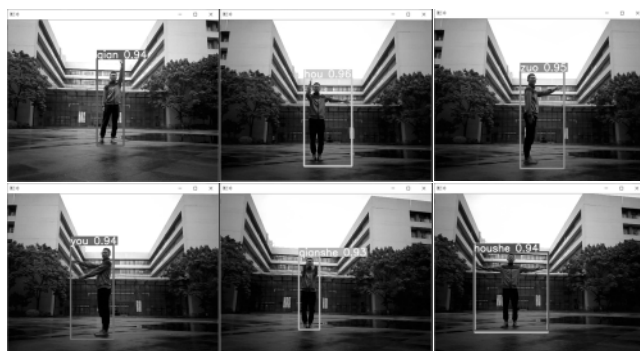
别为:向前96%、94%、94%;向后97%、97%、96%;向左96%、96%、95%;向右96%、95%、94%;前射95%、94%、93%;后射98%、96%、94%,每种姿态识别效果如图5所示。识别准确率如图6所示。



(a)室内识别效果



(b)室外晴天识别效果



(c)室外阴天识别效果

图5 视觉识别效果

其中,在室内、室外晴天、室外阴天3种不同环境下6种姿态的识别结果表明:“向后”的识别率最高,平均准确率为96.3%;“前射”的识别率最低,平均准确率为94%。不同姿态之间识别率存在差异的原因在于每种姿态特征的复杂程度,如对于“前射”姿态较为对称,经过翻转处理后的图片与原图差异不大,数据增强的效果相较“向后”非对称的姿态较差。而由于不同环境光线变化影响,室内平均识别率比室外高,而室外晴天相对室外阴天识别率高。



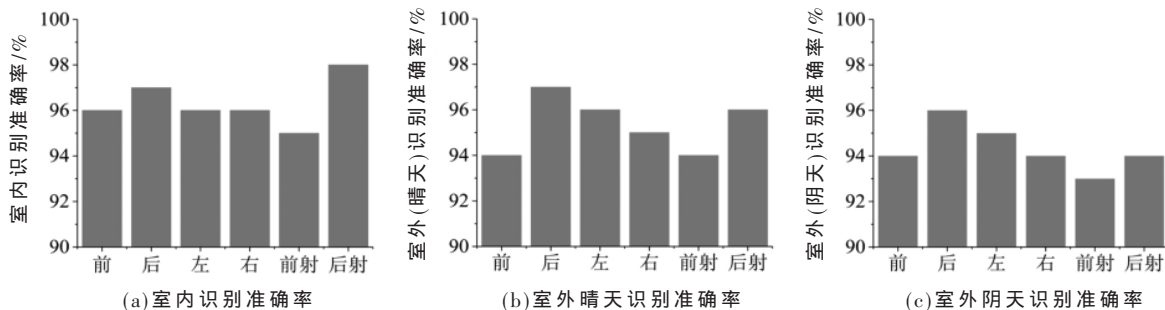


图 6 识别准确率

## 4.2 机器人应用效果

机器人基于视觉的控制效果如图 7 所示,以图中坐标轴为基准,当摄像头捕捉到运动员“向前”姿态时,机器人完成了向前运动;在识别到“后射”姿态时,机器人完成了抛球动作。机器人每当接收到一种新的指令时,机器人会立刻执行相应动作并一直保持,直到下一指令的接收。



(a) 机器人向前运动

(b) 机器人后射运动

图 7 机器人应用效果

## 5 结论

本文提出了一种以树莓派为核心、基于 YOLOv5 视觉检测算法的姿态识别方法,并将其应用于一种橄榄球陪练机器人上。对 3 种不同环境下 6 种姿态的视觉识别实验进行了分析与讨论,结果表明:非对称姿态的识别率较高;光线变化小的室内环境较室外环境识别率高,晴天较阴天识别率高。本实验制作的具有较高识别率的姿态数据集和对不同环境下的实验对比分析为基于视觉的人机交互提供了实验和数据参考。在本实验中,没有进行运动员快速移动过程中不同姿态指令的识别与机器人动作实现,这将是本研究下一阶段的主要方向。

### 参考文献

- [1] 张先强.乒乓球机器人的智能控制方法——击球轨迹的生成[D].北京:北京交通大学,2012.
- [2] 李小康,赵险峰,胡述祥,等.智能乒乓球陪练机器人结

构设计[J].机械设计与制造工程,2020,49(11):49-52.

- [3] 谢小英,曹伟.一种羽毛球陪练机器人的结构设计[J].机械设计与制造工程,2019,48(11):43-47.
- [4] 刘煜,杨平,段丙涛.基于惯性传感器的手机手势识别[J].电子技术应用,2011,37(12):76-79.
- [5] 郭霞,谭亚丽,申森.基于 FDC2214 的手势识别系统[J].传感器与微系统,2018,37(12):95-97.
- [6] 胡青松,张亮.人体行为序列化识别算法研究[J].电子技术应用,2018,44(4):122-125.
- [7] 黄友文,万超伦.基于深度学习的人体行为识别算法[J].电子技术应用,2018,44(10):1-5.
- [8] 王林,黄三丽.YOLOv3 网络在车标检测中的应用[J].电子技术应用,2020,46(9):34-37.
- [9] 江帆,刘辉,王彬,等.基于 CNN-GRNN 模型的图像识别[J].计算机工程,2017,43(4):257-262.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of CVPR 2015, 2015: 779-788.
- [11] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517-6525.
- [12] REDMON J, FARHADI A. YOLOv3: an incremental improvement[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv: 2004.10934v1[cs.CV], 2020.
- [14] 谈世磊,别雄波,卢功林,等.基于 YOLOv5 网络模型的人员口罩佩戴实时检测[J].激光杂志,2021,42(2):147-150.

(收稿日期:2021-05-10)

### 作者简介:

曾杨吉(1996-),男,硕士研究生,主要研究方向:机器人技术、深度学习。

刘自红(1984-),通信作者,男,博士,讲师,主要研究方向:机器人技术、机电一体化,E-mail:270820424@qq.com。

蔡勇(1962-),男,博士,教授,主要研究方向:机器人技术。



扫码下载电子文档

## 版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所