

DOI:10.19651/j.cnki.emt.2518980

# 基于深度学习的自动拾取网球机器人\*

梁宏宇<sup>1,2</sup> 闫坤<sup>1,2</sup> 郝航勃<sup>1,2</sup>

(1. 桂林电子科技大学信息与通信学院 桂林 541004;

2. 桂林电子科技大学卫星导航定位与位置服务国家地方联合工程研究中心 桂林 541004)

**摘要:** 本文提出了一种基于深度学习的自动拾取网球机器人,旨在解决网球场人工捡球效率低、耗时长的问题。该机器人通过集成树莓派 5B 主控板、STM32RCT6 单片机、USB 摄像头和无刷直流电机等硬件模块,结合轻量化的 YOLOv11、改进的 DBSCAN 聚类路径规划算法、双闭环 PID 算法以及碾压滚轴式的捡球机械结构,实现了对网球的高效识别、路径优化和自主拾取。研究中对 YOLOv11 进行了轻量化改进,采用 StarNet 主干网络、C3k2\_Faster 模块和共享卷积轻量化检测头 SCLD-Head,显著降低了模型的计算需求。实验结果表明,改进后的模型参数量减少了 80.8%,GFLOPs 仅为 1.7,mAP@0.75 达到 0.980 6,检测速度达到 129.7 fps。此外,基于 DBSCAN 聚类算法的路径规划通过密度聚类和距离加权模型优化了捡球路径,提高了机器人在复杂环境下的适应性和鲁棒性。实验结果表明,该模型成功部署于树莓派平台后,系统能够在不同光照条件下准确识别网球,并达到 9~12 fps 的检测速度,单次可拾取 7~9 个网球,能够显著提升网球拾取效率,具有重要的应用价值和市场潜力。

**关键词:** 网球拾取机器人;树莓派;YOLOv11 轻量化;路径规划;双闭环 PID;网球拾取结构

**中图分类号:** TP242;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 510.8050

## Automatic tennis ball collection robot based on deep learning

Liang Hongyu<sup>1,2</sup> Yan Kun<sup>1,2</sup> Hao Hangbo<sup>1,2</sup>

(1. School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China;

2. National and Local Joint Engineering Research Center for Satellite Navigation, Positioning and Location Services,

Guilin University of Electronic Technology, Guilin 541004, China)

**Abstract:** This paper presents a deep learning-based autonomous tennis ball retrieval robot designed to address the inefficiencies of manual ball collection. The robot integrates a Raspberry Pi 5B, STM32RCT6 microcontroller, USB camera, and brushless DC motors. Combining a lightweight YOLOv11, an improved DBSCAN clustering-based path planning algorithm, a dual-loop PID controller, and a roller-based collection mechanism, the robot achieves efficient tennis ball recognition, optimized path planning, and autonomous retrieval. The YOLOv11 model was lightened using a StarNet backbone, C3k2\_Faster module, and shared convolutional lightweight detection head, significantly reducing computational demands. Experimental results show an 80.8% reduction in parameters, a GFLOPs of only 1.7, an mAP@0.75 of 0.980 6, and a detection speed of 129.7 fps. The DBSCAN-based path planning, optimized through density clustering and a distance-weighted model, enhances the robot's adaptability and robustness in complex environments. Deployed on a Raspberry Pi, the system accurately recognizes tennis balls under varying lighting conditions, achieves a detection speed of 9~12 fps, and retrieves 7~9 balls per run, demonstrating significantly improved retrieval efficiency and promising practical applications.

**Keywords:** tennis picking robot; Raspberry Pi; light-weighted YOLOv11; path planning; dual closed-loop PID; tennis picking structure

## 0 引言

随着国家现代化建设的稳步推进和人民生活水平的普

遍提高,网球运动逐渐得到了很多人的喜爱。据统计,中国网球产业规模已达 369 亿元,同比增长 15%,网球运动人口达 2 381 万人,占全球 25.3%,这些数据反映出网球项目

收稿日期:2025-05-30

\* 基金项目:国家自然科学基金(62101147)、广西自然科学基金(桂科 2020GXNSFAA159146)、广西创新驱动发展专项(桂科 AA21077008)、教育部重点实验室基金(CRKL190108)项目资助

从初期的“引发关注”阶段发展到“融入生活”的常态化阶段<sup>[1]</sup>。然而,网球运动普及的同时,重复捡球成为一项耗时乏味且影响体验的体力消耗。因此,研发具备自主能力的网球拾取机器人,不仅符合行业发展趋势,且具有显著的市场潜力和应用价值。

近年来,利用智能机器人技术实现网球自动拾取已成为研究热点。然而,在实际应用中,目标对象识别不准确、定位误差以及控制不稳定等问题,直接影响了网球拾取的效率。随着图像处理技术的快速发展,为网球拾取机器人的设计提供了新的思路,通过视觉反馈信息可间接控制机器人的运动,使其能够根据目标对象的颜色、位置等信息做出准确决策,进而驱动机器人末端执行器实现精确运动。文献[2-5]设计了一种基于计算机视觉技术的自主捡球机器人,均通过对图像进行灰度化、二值化和特征提取等操作,采用几何特征和模板匹配法进行多目标球类物体识别并表现出较高的成功率,初步实现了依赖自身视觉系统完成目标识别、导航和自主捡球。然而,这些传统方法主要依赖于颜色和轮廓特征,其对环境变化的适应性不足。聚集的网球和光照条件的变化均会对图像识别过程产生干扰,导致识别精度降低,这些缺陷限制了传统方法在实际复杂场景下的应用,尤其是在环境条件差异较大的情况下,需频繁调整图像处理参数才能保证自动捡球系统的稳定运行,从而增加了系统部署的难度和成本。

随着深度学习的快速发展,基于深度学习的目标识别算法在目标检测中越来越重要。张崢<sup>[6]</sup>研究了智能捡网球机器人的目标识别和路径规划技术,利用卷积神经网络的识别方法,其准确度显著优于传统识别方法。施金汝等<sup>[7]</sup>设计了一种基于 YOLOv5 算法和树莓派的检测系统,采用 YOLO 算法实现对目标的识别,该方法在实时性和准确性方面表现出色,尤其是在处理复杂背景和不同光照下的图像时。此类研究虽提升了检测精度,但在模型的轻量化程度上关注较少,难以部署在储存能力和算力较差的边缘设备中,在实际生产中的泛化性较差。

Ultralytics 发布的最新 YOLO 系列在图像检测领域表现优异,特别适合针对目标检测进行轻量化改进。在轻量化模型研究方面,文献[8-9]通过替换主干网络、检测头和颈部网络模块,显著降低了模型参数数量和计算量,在保持高精度的同时提升了检测速度。类似地,徐尽达等<sup>[10]</sup>提出了 LS-YOLO 模型,该模型采用轻量级网络结构和剪枝技术,不仅减少了参数数量和计算量,还提高了检测性能和鲁棒性。

针对现有网球拾取机器人所面临的挑战,本文提出了一种基于深度学习的自主网球拾取机器人设计方案。该方案以树莓派作为中央控制单元,协同 STM32F103RCT6 微控制器以及多种传感器模块,构建了完整的识别与控制系统。在目标识别方面,采用了轻量化的 YOLOv11 算法,实现对网球的精准且高效的识别;在路径规划方面,对传统

DBSCAN 聚类算法进行了改进,使其更适用于网球拾取场景,优化拾取路径并提升捡球效率;在运动控制方面,采用双闭环 PID 控制算法,以实现机器人的稳定、精确运动。此外,本研究设计的滚轴碾压式拾取机构能够有效地将网球卷入收球框,并具备同时拾取多个网球的能力,从而显著提升整体拾取效率。该方案旨在克服现有技术的局限性,提高网球拾取机器人智能化水平和实用性。

## 1 捡球机器人系统整体设计

### 1.1 总体方案

总体方案设计如图 1 所示,自主拾取网球机器人主要由视觉系统、运动控制系统以及机械结构组成。视觉系统装载在上位机中,主要负责深度学习算法的运算,完成网球的识别并规划捡球路径,利用串口通讯将运动指令发送给下位机;运动控制系统装载于下位机,可根据接收到的上位机的指令以及传感器的信息,从而控制机器人拾取网球。

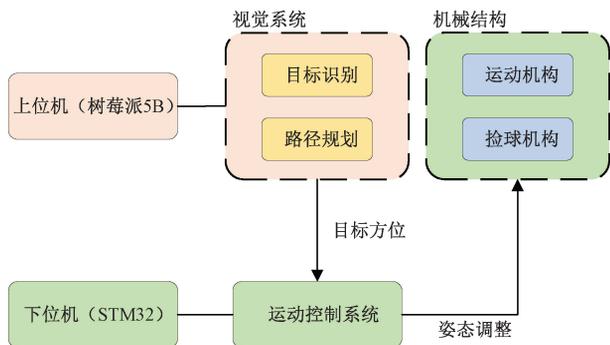


图 1 总体方案设计

Fig. 1 Overall scheme design

### 1.2 网球拾取结构

该碾压式滚轴收球结构主要由塑实材料制成的滚轴和半弧形铝合金挡板构成。滚轴长度约为 60 cm,直径约为 10 cm,作为核心拾球部件。半弧形铝合金挡板长度约为 65 cm,周长约为 20 cm,与滚轴之间保持约 6.45 cm 的间距,形成收球通道。网球直径在 6.35~6.67 cm 之间,由弹性橡胶制成。工作时,网球在滚轴和挡板之间产生弹性形变,被卷入收球通道。随着滚轴持续旋转,网球被推送至收球框,实现批量收球,一次性最多可拾取约 7~9 个网球,具有较高的拾取效率,整体如图 2 所示。

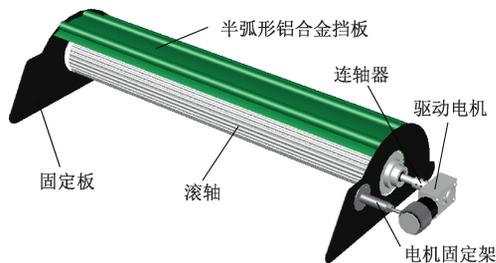


图 2 收球机构

Fig. 2 Ball collection mechanism

## 2 捡球机器人系统硬件设计

### 2.1 图像处理模块

上位机部分采用了树莓派 5B 作为核心计算平台,负责图像处理任务<sup>[11-14]</sup>。树莓派 5B 搭载 2.4 GHz 四核 Cortex-A76 处理器,并配备丰富的硬件接口,具备完善的软件生态系统和强大的 GPU 加速能力,这使其能够高效执行如 YOLOv11 这样的深度学习目标检测算法,满足实时图像处理需求。此外,相较于其他同等性能的嵌入式计算平台,树莓派 5B 具有较高的性价比,为机器人的高效、稳定运行提供了坚实的硬件基础。

### 2.2 主控系统 MCU

控制机器人运动的下位机选择 STM32F103RCT6,其优点有:

- 1) 强大性能: Cortex-M3 内核、72 MHz 主频和硬件加速,满足网球机实时控制需求。
- 2) 丰富外设: 多路 16 位 PWM、12 位 ADC、多个定时器和 UART/SPI/I2C 接口,便于电机精确控制、传感器高精度采集和可靠通信。
- 3) 高性价比: 性能、外设、价格平衡,批量采购更具优势,售价仅十多元。
- 4) 封装体积小: 本身小巧轻便,其 LQFP 封装尺寸约为  $10\text{ mm}^2$ ,重量不足  $0.5\text{ g}$ ,在系统中容易集成。

## 3 捡球机器人总体软件及各算法设计

### 3.1 总体软件架构设计

捡球机器人软件系统采用分层架构,主要包含上位机软件和下位机软件两部分。上位机主要采用了轻量化 YOLOv11 目标检测算法以及基于 DBSCAN 聚类的路径规划算法,下位机软件则负责运行运动控制算法并执行相应的控制指令。同时,为实现上下位机间的数据交互,制定了相应的通信协议。

### 3.2 基于改进的 YOLOv11n 的检测网络

#### 1) YOLOv11n 的轻量化设计

YOLOv11 是由 Ultralytics 公司于 2024 年提出的新一代目标检测模型,其架构由骨干网络、颈部网络和检测头 3 部分组成,通过多项创新机制实现了精度与效率的平衡。骨干网络基于改进的 CSPDarknet53 框架,采用 5 次下采样生成多尺度特征图(P1-P5),并引入轻量化 C3k2 模块替代传统 C2f 模块,通过小核卷积优化信息流以降低计算成本;同时集成 SPPF 模块,利用多尺度池化增强小目标特征捕获能力,并结合 C2PSA 机制嵌入 PSA,通过多层次多头注意力提升特征敏感度。颈部网络采用 PAN-FPN 结构,通过自底向上路径融合浅层定位信息与深层语义特征,优化特征金字塔的定位精度。检测头延续了 YOLOv8 的解耦式检测头,分类分支引入深度可分离卷积(depthwise convolution, DWConv)压缩参数量,回归分支结合分布焦

点损失(distribution focal loss, DFL)与 CIoU 损失函数提升边界框预测精度。实验表明, YOLOv11 在保持实时推理性能的同时,显著提升了小目标检测能力,其模块化设计为复杂场景提供了高效、轻量化的解决方案,在计算效率与检测精度上达到先进水平。

为满足网球拾取机器人系统在嵌入式平台部署场景下的实时目标检测需求,本研究针对 YOLOv11n 网络开展轻量化改进。首先,在主干网络部分,引入 StarNet 模块替代原始特征提取单元,该模块通过融合深度可分离卷积与通道注意力机制构建轻量化卷积结构,在保留多尺度特征表达能力的同时缩减参数规模;其次,颈部网络通过采用 C3k2\_Faster 显著减少参数量与计算复杂度。此外,采用共享卷积轻量检测头 SCLD 替代传统检测头,通过共享卷积层减少计算复杂度;最后,针对网球检测任务的单尺度特性定向精简网络,将整体卷积核数量与输出特征通道数压缩至原始的  $1/8$ ,改进后的模型结构如图 3 所示。

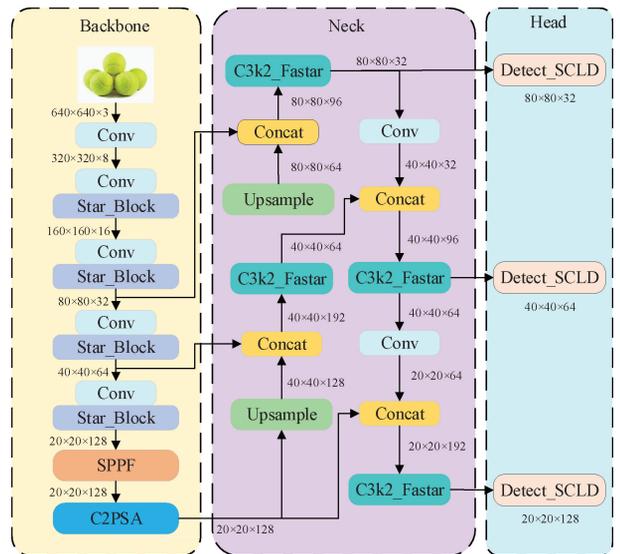


图3 改进后 YOLOv11n 网络结构

Fig. 3 Network structure drawing of improved YOLOv11n

#### 2) StarNet 网络设计

为了提高模型在目标检测任务中的性能,增强特征提取能力,减少计算量和参数量,本文使用 StarNet 代替原有的主干。StarNet 是一种新型的轻量化卷积神经网络架构,其实现基于 DWConv 与动态特征调制的协同设计,通过构建层级的星型拓扑结构实现高效特征提取。网络主体由 stem 层与 4 个渐进式特征转换阶段构成,每个阶段包含下采样模块与串联的星型块(StarBlock)。

如图 4 所示,每个 StarBlock 的核心创新在于双路动态特征交互机制,首先对输入特征执行  $7 \times 7$  DWConv 捕获上下文信息并用 ReLU6 函数激活,随后通过两个并行的  $1 \times 1$  卷积将特征投影至高维空间,对其中一条路径采用 ReLU6 激活函数并与另一路径进行逐元素乘法(即 Star),实现非线性的特征调制。如式(1)所示。

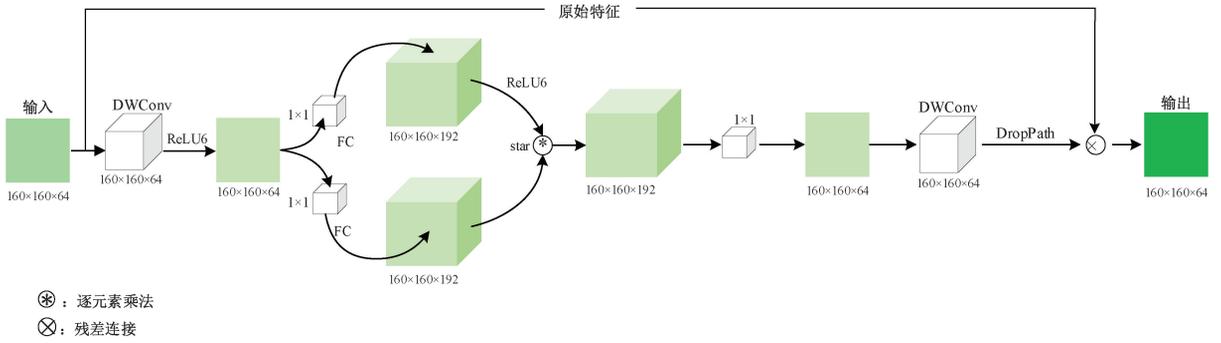


图 4 StarBlock 网络结构

Fig. 4 Network structure drawing of StarBlock

$$\begin{cases} X_{mid} = DWConv(X_{in}) \\ X_{branch} = ReLU6(f_1(X_{mid})) \odot f_2(X_{mid}) \end{cases} \quad (1)$$

其中,  $X_{in}$  表示输入特征图,  $DWConv(\cdot)$  表示深度可分离卷积计算,  $X_{mid}$  表示深度卷积后的中间特征图,  $ReLU6(\cdot)$  表示激活函数计算,  $f_1(\cdot)$  和  $f_2(\cdot)$  表示卷积升维操作,  $\odot$  表示逐元素乘法操作,  $X_{branch}$  表示经过分支路径处理后的中间特征图。

调制后的特征经  $1 \times 1$  卷积降维后, 通过第 2 个  $7 \times 7$  深度卷积强化局部细节。为了增强模型的泛化能力, 引入了  $DropPath(\cdot)$  随机丢弃路径, 通过随机丢弃部分通道来减少过拟合的风险, 从而在保证计算效率的同时, 保持模型的高表达能力和泛化能力。最终与跳跃连接的原输入特征相加, 形成残差学习框架。如式(2)所示。

$$\begin{cases} F(x) = DWConv(g(X_{branch})) \\ X_{out} = X_{in} + DropPath(F(x)) \end{cases} \quad (2)$$

其中,  $g(\cdot)$  表示卷积降维操作,  $F(x)$  表示为残差函数,  $DropPath(\cdot)$  表示随机正则化操作,  $X_{out}$  表示输出特征图。

通过以上步骤, StarNet 实现了全局信息的共享与融合, 提升了特征表达能力, 具有参数量少、全局信息交互和特征增强等优势。因此, 本研究选择 StarNet 作为 YOLOv11 的主干网络, 旨在充分利用 StarNet 的优势, 有效降低模型的计算复杂度, 使其能够满足实时性要求, 并更易于部署在嵌入式平台上。

### 3) C3k2\_Faster 模块

针对模型轻量化需求, 对目标检测颈部网络进行结构化改进, 将原有的 C3k2 模块替换为 C3k2\_Faster 模块, 如图 5 所示。C3k2\_Faster 通过重构传统 C3k2 架构的深层特征提取路径, 将原 Bottleneck 结构替换为基于 FasterNet 思想的 FasterBlock, 在保留双分支并行结构的基础上实现计算效率的显著优化。改进后的模块采用浅层分支维持高分辨率特征的空间细节信息, 深层分支则通过堆叠 FasterBlock 实施选择性特征增强。每个 FasterBlock 采用通道分割策略, 将输入特征按预设比例划分为卷积处理区域( $1/4$  通道)与恒等映射区域( $3/4$  通道), 前者通过  $3 \times 3$

PConv 算子进行空间特征提取, 后者直接保留原始通道信息以抑制冗余计算。处理后的特征经  $1 \times 1$  Conv 层完成通道维度融合并结合 BN-ReLU 激活函数增强非线性表达能力, 最终通过  $1 \times 1$  的 Conv 层后与 PConv 层的输入进行残差拼接得到最终输出。因此, 每个 Faster Block 只需对原来的部分通道进行卷积, 显著降低了模型的计算复杂度, 适用于边缘端实时检测任务。

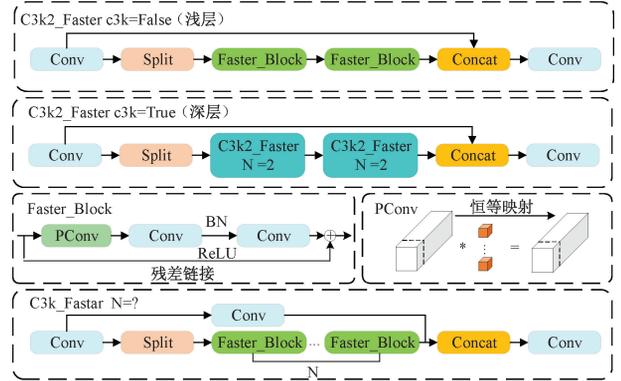


图 5 C3k2\_Faster 模块网络结构

Fig. 5 Network structure drawing of C3k2\_Faster

### 4) SCLD-Head

在 YOLOv11 检测头中, 针对多尺度特征层(如 P3/P4/P5)的每个层级均独立构建分类与回归卷积模块, 导致模型复杂度与计算开销显著增加。为克服上述局限性, 本研究提出一种共享卷积轻量化检测头(shared convolution lightweight detection head, SCLD-Head), 其结构如图 6 所示。首先, 针对输入的多尺度特征层(P3-P5)分别进行  $3 \times 3$  归一化卷积(Conv\_GN)进行空间特征增强; 随后, 将各层级特征输入共享卷积网络(Share\_Conv)进行跨尺度特征融合, 该网络由 DWConv 与逐点卷积级联构成, 通过权重共享机制实现跨尺度特征的统一表征, 同时显著降低计算复杂度; 最终, 各特征层分别通过轻量化分类卷积链(Conv\_Cls)与回归卷积链(Conv\_Reg)生成目标类别概率分布及边界框坐标修正参数。实验表明, SCLD-Head 通过统一化多尺度特征的卷积参数空间, 在维持检测精度的条

件下有效降低模型复杂度,实现计算成本与检测性能的最优平衡。

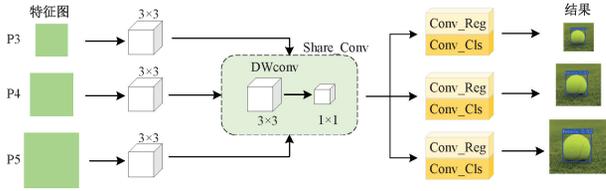


图6 SCLD-Head结构示意图

Fig. 6 SCLD-Head structural diagram

### 3.3 基于改进 DBSCAN 聚类算法的路径规划

本研究采用基于密度的 DBSCAN 算法<sup>[15]</sup>对网球分布进行聚类分析,以解决动态场景下的多目标路径规划问题。该算法的核心在于通过设定邻域半径  $\epsilon$  和最小邻域点数  $MinPts$  两个关键参数,将密度相连的点聚类为同一簇。具体而言,算法将摄像头采集的每一帧图像视为二维坐标系,其中每个检测到的网球中心点  $(x_i, y_i)$  坐标构成数据集  $X = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ ,  $n$  为当前帧的网球数量。对于任意点  $p \in X$ , 其  $\epsilon$  邻域内的点数需满足  $|N_\epsilon(p)| \geq MinPts$  方可标记为核心点,其中  $N_\epsilon(p) = \{q \in X \mid dist(q, p) \leq \epsilon\}$ ,  $dist(q, p)$  为点  $p$  与  $q$  的欧氏距离。算法递归合并核心点的密度可达区域形成簇,而无法归类的点则判定为噪声点,由此实现任意形状簇的划分及噪声过滤。使用 DBSCAN 算法对这些点进行聚类分簇,旨在将空间位置上较为接近的网球归为同一簇,从而为后续的路径规划奠定基础。

为量化各簇的拾取优先级,本研究提出基于距离加权的价值评估模型。针对机器人当前位置  $(x_r, y_r)$ , 计算簇  $C_k$  内每个网球中心点的欧氏距离  $d_i$ 。如式(3)所示。

$$d_i = \sqrt{(x_i - x_r)^2 + (y_i - y_r)^2} \quad (3)$$

并定义每个网球的权重函数  $w_i = 1/(d_i + \delta)$  ( $\delta$  为平滑因子以避免零分母),其物理意义为距离越近的网球权重越高。簇的综合价值由权重总和表征,即  $W_k = \sum_{i \in C_k} w_i$ , 最终选取最大的簇  $W_k$  作为目标簇  $C_i$ 。此策略通过权衡空间密度与邻近性,确保机器人优先处理高价值区域。

目标簇选定后,通过质心坐标  $(x_c, y_c)$  引导全局路径,质心坐标如式(4)、(5)所示。

$$x_c = \frac{1}{C_i} \sum_{i \in C_i} x_i \quad (4)$$

$$y_c = \frac{1}{C_i} \sum_{i \in C_i} y_i \quad (5)$$

同时,机器人运动距离由簇内最远网球决定,以此确保小车拾取路径覆盖目标簇的全部区域。该方法通过 DBSCAN 的密度聚类特性对识别到的网球进行分组,结合距离加权模型动态适配环境变化,驱动小车朝着距离较近且密度较大的最优簇行进,最终实现高效、鲁棒的多目标

拾取路径规划。

### 3.4 双闭环 PID 算法

本研究采用了双闭环 PID 控制策略,旨在实现对目标物体的精确跟踪和距离保持<sup>[16-19]</sup>。横向位置控制环路以图像宽度中心作为期望位置,通过计算目标质心  $x_c$  坐标与中心点的偏移量,得到横向修正量。PID 控制器基于此偏差,输出转向修正系数  $x_{out}$ , 用于调整系统的横向运动,确保目标始终位于视野中心。距离控制环路则以预设安全距离 10 cm 为目标,构建了一个非线性距离估算模型,将簇内最远小球在图像坐标系中的二维像素坐标转换为真实空间距离。PID 控制器基于当前距离与安全距离的偏差,输出速度调节系数  $r_{out}$ , 用于动态调整系统的接近速度,保持目标物体在安全距离范围内。两个 PID 控制器采用了相同的整定参数,其中较大的比例系数 P 保证系统对目标偏移的快速响应;较小的积分项 I 用于消除稳态误差,确保系统最终能够精确达到期望位置;微分量 D 则用于抑制摄像头抖动带来的高频噪声,提高系统的稳定性。在每个视觉检测周期内,系统都会执行 PID 运算,分别计算  $x_{out}$  和  $r_{out}$ 。最后,系统将  $x_{out}$  和  $r_{out}$  通过位姿参数映射,转换为电机 PWM 信号,从而驱动系统运动。这种双闭环 PID 控制设计使得系统既能快速跟踪目标质心的横向偏移,又能根据目标距离动态调整行进速度,形成一个自适应控制体系,实现基于视觉反馈的实时闭环运动控制。

## 4 运动控制

### 4.1 硬件搭建

为满足自动网球收集机器人不同功能模块的电力需求,本系统采用 3 组独立锂电池供电方案。动力系统采用 24 V 锂电池驱动两个无刷霍尔电机轮,利用无刷电机低速高扭矩特性,保证车辆在负载变化、频繁启停等工况下的稳定运行。鉴于树莓派及 STM32RCT6 分别负责 YOLOv11 目标检测和运动控制,计算量大导致功耗较高,因此采用独立电池供电,以避免动力系统干扰,确保控制系统的稳定性和可靠性。收球装置仅需持续旋转,故采用 12 V 锂电池单独供电。本实验采用无刷直流电机,因其高效率、长寿命、精准控制、高扭矩及低维护等优势,满足了自动网球收集小车对动力稳定性、环境适应性与经济性的核心需求,是理想的驱动电机选择。

### 4.2 具体运动过程

网球拾取机器人的整体结构示意图如图 7 所示。系统启动后,电源管理模块负责对各功能单元进行分布式供电。树莓派首先控制摄像头进行实时图像采集,并利用轻量化 YOLOv11 算法进行网球检测。若检测到网球,系统采用改进的 DBSCAN 聚类算法计算当前最优目标簇的质心坐标与距离信息,并通过双闭环 PID 控制算法将这些信息转换为差速驱动轮的 PWM 调速信号,从而控制机器人运动。单片机在完成轮子转速调整后,通过串口向树莓派

发送执行完毕信号。收到该信号后,树莓派对当前图像帧进行重新检测,并再次执行 DBSCAN 聚类算法,对剩余网球进行聚类和目标簇选择。新的质心坐标和前进距离将作为下一个控制周期的双闭环 PID 目标设定值,驱动机器人前往下一个目标簇。为应对运动过程中网球分布的动态变化,系统在每个视觉检测周期内重新执行聚类算法,确保路径规划的实时性和自适应性。

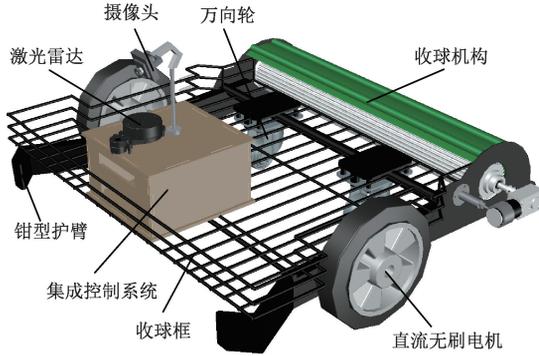


图 7 网球及整体结构

Fig. 7 Tennis racket and overall structure

针对视觉目标丢失情况,考虑到目标检测模型在树莓派平台上可实现 9~12 fps 的运行速度,因此,当连续 20 帧图像中均未能检测到网球目标时,系统将触发广角搜索机制。具体而言,机器人将原地旋转,分步扫描周围环境。由于摄像头广角约为  $50^\circ$ ,为确保完整覆盖  $360^\circ$  范围,所以机器人将进行至少 8 次旋转,以实现全方位扫描搜索目标。完成单次旋转后,系统将再次连续检测 20 帧图像,判定是否识别到网球,若重新识别到网球,则立即启动捡球程序。若小车累计旋转 8 次后仍未能发现目标,系统将进入关机状态。具体的运动控制流程如图 8 所示。同时,在机器人的运动过程中,滚轴捡球机构由电机驱动并持续转动。当网球触碰到滚轴时,将被碾入收球通道并输送至收球框内,从而完成网球的收集过程。

## 5 实验与结果分析

### 5.1 构建数据集

网球数据集包含 500 张网球图片,图像采集自不同角度和背景,其中 200 张来源于公开的百度飞桨网球图像数据集,另有 300 张为自行采集的网球场散落网球图片以及各种背景的网球图片。数据集按 7:1.5:1.5 的比例划分为训练集、验证集和测试集。

为避免深度学习模型训练过程中的过拟合现象,并获得更优的泛化性能,本研究采用了数据增强策略。对原始数据集图像实施了包括位移、错切、添加噪声以及色域扭曲等多种变换,将训练集扩充至 3 800 张。

### 5.2 测试环境构建

实验环境统一配置如下:处理器为 Intel<sup>®</sup> Core™ i5-12600KF,运行内存 16 GB,GPU 为 GeForce RTX 4060

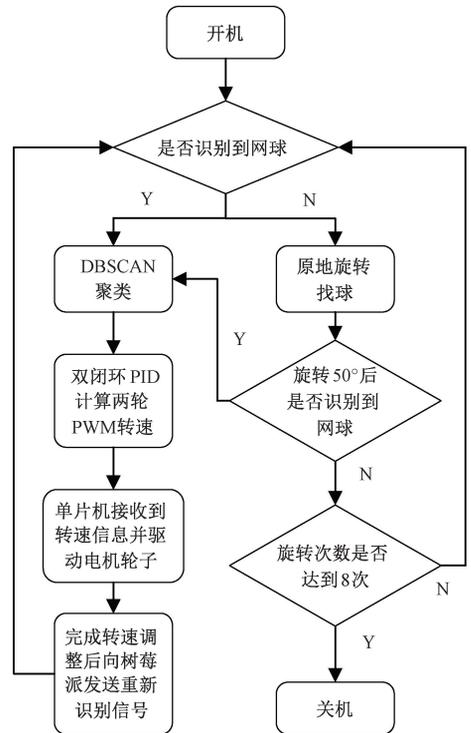


图 8 运动控制流程图

Fig. 8 Motion control flowchart

(8 GB),操作系统为 Windows11。使用 PyTorch 1.13 和 CUDA12.6 版本进行加速。训练过程遵循统一超参数,共进行 200 轮。

### 5.3 网球识别算法性能

针对识别算法的性能,本研究从检测精度和轻量化性能两方面进行了综合评估。在检测精度方面,利用模型平均检测精度  $mAP@0.75$  评估。在轻量化性能方面,通过参数量衡量模型规模,用浮点运算数 GFLOPs 表示计算复杂度,每秒帧数 FPS 评估检测速度。为验证所提出改进方案的有效性,本研究以 YOLOv11n 为基准模型,在保持软硬件环境一致的前提下,通过消融实验来检验轻量化改进效果,结果如表 1 所示。由表 1 可知,针对网球目标检测,本研究改进的每个阶段相较基准模型均有提升,模型 B 引入 Starnet 主干网络后模型参数量、GFLOPs 和  $mAP@0.75$  分别下降约 26.9%、21.9%、0.13%,FPS 提升 14.1 fps;模型 C 引入 C3k2-Faster 模块后参数量、GFLOPs 和  $mAP@0.75$  分别下降约 11.5%、7.8%、0.54%,FPS 提升 10.7 fps;模型 D 引入 LSCD-Head 检测头后参数量、GFLOPs 和  $mAP@0.75$  分别下降约 7.7%、12.5%、0.41%,FPS 提升 12.3 fps;模型 E~H 进一步验证改进模块的有效性,并分析不同模块组合对性能的影响。实验结果表明,尽管各模块在网络结构中的位置及对检测性能的影响有所差异,但各项改进措施协同作用能够有效提升模型的整体检测性能。具体而言,模型参数量与浮点运算数呈现出累积式下降趋势,充分验证了本研究的轻量化设计目

标。通过进一步缩减模型 H 的通道数至 1/8, 尽管 mAP@0.75 略微下降 0.36%, 却实现了参数量和 GFLOPs 分别下降约 61.5% 和 46.9%, 同时 FPS 提升了 16.3 fps。最终, 相较于基线模型, 改进后的模型检测速度提升了 39.6 fps, mAP@0.75、参数量和 GFLOPs 分别降低了 0.76%、80.8% 和 73.4%。虽 mAP@0.75 略有下降, 但考虑其大于 95% 时模型已具备高定位准确性, 且轻量化导致

的精度损失主要体现为边界框的微小偏差, 而非显著的漏检或误检, 尤其考虑到网球的显著性和非孤立分布特征, 以及机器人可批量拾取的特性, 轻微的边界框偏移对实际拾取成功率的影响甚微。综上, 本研究提出的轻量化策略在保证可接受的检测精度前提下, 显著降低了模型计算复杂度, 提升了检测速度, 为嵌入式部署奠定了基础。

表 1 消融实验

Table 1 Ablation experiment

模型 编号	Starnet	C3k2-Faster	SCLD-Head	精简网络	目标检测网络			
					mAP@0.75	参数量/M	GFLOPs	FPS/fps
A					<b>0.988 1</b>	2.6	6.4	90.1
B	✓				0.986 8	1.9	5.0	104.2
C		✓			0.982 8	2.3	5.9	100.2
D			✓		0.984 0	2.4	5.6	102.4
E	✓	✓			0.985 2	1.8	4.8	109.6
F	✓		✓		0.980 7	1.8	4.3	105.8
G		✓	✓		0.978 4	2.1	5.7	104.4
H	✓	✓	✓		0.984 1	1.3	3.2	113.4
I	✓	✓	✓	✓	0.980 6	<b>0.5</b>	<b>1.7</b>	<b>129.7</b>

为验证本研究算法的轻量化优势, 本文将最终模型与当前主流的目标检测模型, 包括 SSD、YOLO 系列模型以及 YOLO 系列模型进行了对比实验。实验结果汇总于表 2, 其中最优性能指标以加粗字体显示。针对网球检测任务, SSD-VGG16 展现出最高的检测精度, 但检测速度方面表现不足; 虽然采用轻量化 MobileNetv2 改进后参数量大幅度减少, 但是检测速度却有所下降。综合考虑精度与轻量化, YOLOv11n 在前 8 个经典模型中实现了最佳的平衡。相较于 YOLOv11n, 本文提出的最终模型在检测精度

仅下降 0.76% 的情况下, 显著降低了参数量、模型计算量和模型规模, 并提升了检测速度。为满足嵌入式部署需求, 本研究将训练后的 PyTorch 模型(.pt)转换为 NCNN 格式, 并在树莓派平台部署。NCNN 框架凭借其 ARM 架构的深度优化, 能够有效降低内存占用和推理耗时, 同时具备实时性、低功耗和跨平台兼容等优势, 使其成为嵌入式场景下的理想推理框架。实验数据表明, 采用 NCNN 框架后, 模型在树莓派平台上可实现 9~12 fps 的推理速度, 满足实时检测应用的需求。

表 2 模型对比实验

Table 2 Contrast experiment

模型	目标检测网络				
	mAP@0.75	参数量/M	GFLOPs	模型大小/M	FPS/fps
SSD-VGG16	<b>0.993 5</b>	26.3	62.7	90.6	72.8
SSD-MobileNetV2	0.991 2	6.2	1.8	14.3	67.3
YOLOv7-tiny	0.982 5	6.0	13.2	12.0	116.6
YOLOv8n	0.988 2	3.0	8.2	6.1	102.3
YOLOv9-tiny	0.987 4	2.0	7.8	16.9	91.5
YOLOv10n	0.986 1	2.7	8.4	5.8	99.2
YOLOv11n	0.988 1	2.6	6.4	5.3	90.1
YOLOv11s	0.987 1	9.4	21.5	19.2	90.6
本文	0.980 6	<b>0.5</b>	<b>1.7</b>	<b>1.2</b>	<b>129.7</b>

#### 5.4 捡球机器人整体工作性

为了验证网球拾取机器人是否具备高效的目标识别、

路径规划能力以及精确的运动控制性能, 本研究于标准网球场地的半场区域开展了针对性实验。为考察不同因素对

系统性能的影响,实验设计了多组参数组合,包括网球数量、光照条件以及场景类型。针对每组参数组合,均进行重复测试取平均值。本研究将场景类型分为聚集型和分散型两种。聚集型场景被定义为有多处球体聚集密度 $\geq 8$ 个/ $m^2$ 且球体间距 $\leq 0.3$  m,该指标的设置旨在模拟实际应用中常见的球体堆积情况。聚集型场景如图 9 左图所示,分散型场景如图 9 右图所示。实验启动阶段,通过无线 Wi-Fi 局域网建立 PC 端与树莓派之间的 VNC 连接,以便于在 PC 端实时监测图像识别结果,如图 10 所示。

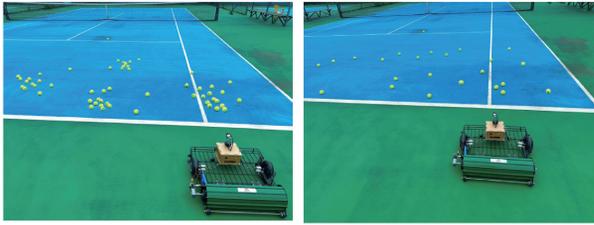


图 9 实验场景

Fig. 9 Experimental setup

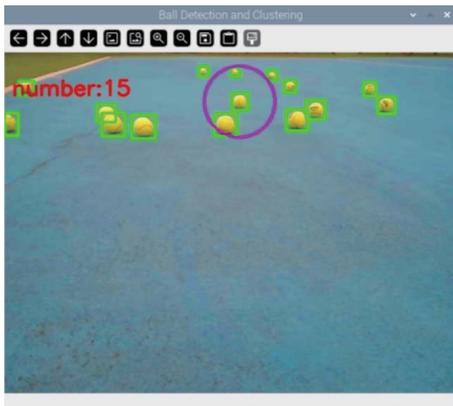


图 10 上位机实时检测界面

Fig. 10 Real-time monitoring interface of the host computer

为了系统评估网球识别精度和路径规划效率,本研究设计了两组独立的实验。网球识别精度的评估实验设计如表 3 所示,识别精度通过计算正确识别的网球数量占总网球数的百分比来量化。路径规划效率评估实验设计如表 4 所示。为了量化路径规划效率,本实验采用固定时间窗口(1 min)的方式,以单位时间内成功拾取的球体数量作为主要评估指标。具体而言,记录机器人在一次完整拾取任务中每分钟成功拾取的网球数量,以此评估拾取效率。为减少机械及环境因素对实验结果的干扰,确保实验条件的一致性,在每次完整拾取任务中均尽可能排除此类影响。随后,将基于 DBSCAN 聚类算法的路径规划方法(下文简称“本文方法”)与无聚类就近拾取策略进行对比分析。实验对比指标包括:单位时间拾取球体数量、机器人完成网球拾取任务的总耗时,以及拾取成功率。为更清晰地界定评估指标,本文将总耗时定义为从机器人开始进

行网球识别到关机所经历的时间,拾取成功率定义为成功拾取的网球数量占环境中总网球数量的百分比。所有指标均通过多次重复实验进行测量,并计算平均值,以降低随机误差,提高实验结果的统计可靠性。

表 3 识别精度测试结果统计

实验条件	场景类型	球个数	识别率/%
光照充足	分散	1~20	100
		21~40	98.7
	聚集	1~20	100
		21~40	92.5
光照不足	分散	1~20	100
		21~40	95.3
	聚集	1~20	97.6
		21~40	87.5

表 4 路径规划效率测试结果统计

实验方式	场景类型	球个数	单位时间内拾取球数	总耗时/s	拾取成功率/%
本文方法	分散	1~20	10.4	115.8	94.5
		21~40	11.6	212.5	87.6
	聚集	1~20	12.7	88.2	98.2
		21~40	16.2	150.5	93.7
就近拾取	分散	1~20	8.2	131.6	89.6
		21~40	9.3	232.1	76.4
	聚集	1~20	10.4	110.4	91.3
		21~40	12.6	213.7	85.8

如表 3 所示,该算法在光照充足、球体分散且数量较少的情况下表现最佳,即便在球体数量增加、场景聚集或光照不足等不利条件下,识别准确率仍保持较高水平,验证了算法的有效性鲁棒性,为后续的路径规划奠定基础。

如表 4 所示,在分散场景下,“本文方法”在单位时间内拾取球体数量上略高于“就近拾取”,但耗时更短;在聚集场景下,无论球数多少,“本文方法”在单位时间内拾取球体数量均显著优于“就近拾取”,同时耗时更短。上述结果表明,在分散和聚集两种场景下,“本文方法”在单位时间内拾取球体数量上均优于“就近拾取”。这归因于 DBSCAN 聚类算法能够对球体进行分组,从而引导机器人规划出更优化的拾取路径,减少了不必要的转向和行驶距离。同时,在两种场景下的拾取成功率均表现良好,这得益于机器人拾取网球覆盖面积较大和拾取结构的可靠性。然而,实验中仍存在少数拾取失败的情况,这主要源于机器人快速运动或光照变化剧烈时,视觉系统可能暂时无法准确识别远处球体,导致机器人未能到达拾取位置,忽略

该区域的网球。相比之下,由于“就近拾取”易陷入局部最优,就会导致拾取成功率相对较低。例如,在有多个球体簇的场景中,若机器人仅选择邻近的单个球体拾取,可能导致其偏离包含更多球体的潜在拾取区域,该策略在短期内看似高效,但在整个拾取任务中会增加机器人行驶总距离,产生额外的转向操作,从而增加更多识别误差的风险,降低整体拾取成功率。综合来看,该网球拾取机器人在各种测试条件下均表现出良好的性能,验证了其在目标识别、路径规划和网球拾取方面的有效性。

## 6 结 论

本文提出了一种基于树莓派深度学习的自动拾取网球机器人,通过集成树莓派 5B 主控板、STM32RCT6 单片机、USB 摄像头和无刷直流电机等硬件模块,结合轻量化的 YOLOv11 目标检测算法、改进的 DBSCAN 聚类路径规划算法以及碾压滚轴式的拾取机械结构,成功实现了对网球的高效识别、路径优化和自主拾取。该机器人系统设计完整,视觉系统基于树莓派 5B 负责目标检测和路径规划,运动控制系统基于 STM32RCT6 负责执行运动控制指令,上下位机通过串口通信完成数据交互,确保了系统的高效运行。

实验结果表明,在目标检测方面,改进后的模型参数量减少了 80.8%,GFLOPs 仅为 1.7,mAP@0.75 达到 0.9806,检测速度达到 129.7 fps,有效满足了嵌入式平台的实时性要求。在路径规划方面,算法能够有效规划拾取路径,确保机器人优先处理高价值区域,从而提高拾取效率。在运动控制方面,采用双闭环 PID 控制策略,实现了对目标物体的精确跟踪和距离保持,确保机器人能够稳定、精确地运动。该机器人能够在不同光照条件下准确识别网球,检测速度达到 9~12 fps,单次可拾取 7~9 个网球,显著提升了拾取效率。未来的研究将进一步优化算法和硬件设计,提升机器人的智能化水平和适应性,拓展其在其他领域的应用。

## 参考文献

[1] 王园园. 全民健身背景下大众网球运动的现状及发展对策研究[J]. 当代体育科技, 2024, 14(35):92-95.  
WANG Y Y. Research on the current situation and development countermeasures of mass tennis under the background of national fitness [J]. Contemporary Sports Technology, 2024, 14(35):92-95.

[2] 邓世春, 潘佳松, 邓君, 等. 基于树莓派和 OpenCV 视觉技术的自动捡网球机器人[J]. 机电工程技术, 2024, 53(7):60-63.  
DENG SH CH, PAN J S, DENG J, et al. Automatic tennis ball picking robot based on raspberry pi and OpenCV vision technology[J]. Journal of Mechanical and Electrical Engineering Technology, 2024, 53(7):60-63.

[3] CHEN Y. Intelligent tennis ball pick-up trolley based on visual recognition [C]. 2025 International Conference on Electrical Drives, Power Electronics & Engineering(EDPEE). IEEE, 2025: 628-633.

[4] 彭磊. 视觉导航的自主捡球机器人设计与开发[D]. 南京:东南大学, 2019.  
PENG L. Design and development of autonomous ball-picking robot based on visual navigation[D]. Nanjing: Southeast University, 2019.

[5] SHEN C, SUN Z. Research on target localization recognition of automatic mobile ball-picking robot[J]. Journal of Optics, 2022, 51(4): 866-873.

[6] 张峥. 智能捡网球机器人的目标识别定位和路径规划研究[D]. 北京:北京林业大学, 2020.  
ZHANG ZH. Research on target recognition, positioning and path planning of intelligent tennis ball picking robot [D]. Beijing: Beijing Forestry University, 2020.

[7] 施金汝, 许凌峰, 陈金立, 等. 基于树莓派和 YOLOv5 算法的路面裂缝检测系统[J]. 信息技术, 2022,(4):22-28,34.  
SHI J R, XU L F, CHEN J L, et al. Road crack detection system based on raspberry pi and YOLOv5 algorithm[J]. Information Technology, 2022(4):22-28,34.

[8] 方明, 张娇, 徐晶, 等. 基于改进 YOLOv8 的轻量化皮革缺陷检测方法[J]. 电子测量技术, 2025, 48(1): 111-118.  
FANG M, ZHANG J, XU J, et al. A lightweight leather defect detection method based on improved YOLOv8[J]. Electronic Measurement Technology, 2025, 48(1):111-118.

[9] 刘忠英, 翟鹏飞, 侯维岩. 基于嵌入式平台和轻量化模型的板材计数装置[J]. 电子测量技术, 2024, 47(9):46-51.  
LIU ZH Y, ZHAI P F, HOU W Y. A plate counting device based on embedded platform and lightweight model [J]. Electronic Measurement Technology, 2024, 47(9):46-51.

[10] 徐尽达, 陈慈发, 张上. 基于轻量级算法的水上垃圾小目标检测研究[J]. 电子测量技术, 2024, 47(18): 145-154.  
XU J D, CHEN C F, ZHANG SH. Research on small target detection of floating garbage based on lightweight algorithms [J]. Electronic Measurement Technology, 2024, 47(18):145-154.

[11] CHEN X, FAN C, YAO W H. Underwater target detection and embedded deployment based on lightweight

- YOLO\_GN [J]. Journal of Supercomputing, 2024, 80(10):14057-14084.
- [12] 许志杰, 吴黎明, 张巧芬, 等. 高性能实时轻量化嵌入式缺陷检测网络的构建[J]. 电子测量与仪器学报, 2025, 39(4):193-202.
- XU ZH J, WU L M, ZHANG Q F, et al. Construction of a high-performance, real-time, lightweight embedded defect detection network [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(4):193-202.
- [13] 陈浩安, 李晖, 黄瑞, 等. 基于树莓派 4B 的无人机动态追踪平台设计[J]. 电子测量技术, 2024, 47(6):182-189.
- CHEN H AN, LI H, HUANG R, et al. Design of a drone dynamic tracking platform based on raspberry pi 4B [J]. Electronic Measurement Technology, 2024, 47(6):182-189.
- [14] TAHILRAMANI N, AHIR P, SAXENA S, et al. Edge-based AI solution for enhancing urban safety: Helmet compliance monitoring with YOLOv9 on raspberry pi [J]. Discover Internet of Things, 2025, 5(1):25-25.
- [15] KAZEMI U, SOLEIMANI S. A new approach data processing: Density-based spatial clustering of applications with noise (DBSCAN) clustering using game-theory [J]. Soft Computing, 2025, 29(3):1-16.
- [16] CHEN X, SUN J. Revising the motion control parameter optimization research of a two-wheel differential car [J]. Actuators, 2024, 13(12):504-504.
- [17] JI D, LI H, LIU W. Methodology and research of intelligent new energy vehicle motion control system based on fuzzy adaptive [J]. International Journal of Vehicle Structures & Systems, 2023, 15(3):372-379.
- [18] 周正阳, 潘树国, 蔚保国, 等. 基于 ABMSSA 的 PP 算法分布式自动驾驶轨迹跟踪控制策略 [J]. 电子测量与仪器学报, 2024, 38(6):50-57.
- ZHOU ZH Y, PAN SH G, WEI B G, et al. Distributed autonomous driving trajectory tracking control strategy of pp algorithm based on ABMSSA [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(6):50-57.
- [19] JIAN M, ZHU S, DU W, et al. Design of a temperature control system for transient thermal measurement utilizing the backpropagation PID algorithm [J]. International Journal of Thermal Sciences, 2025, 210:109654.

### 作者简介

梁宏宇, 硕士研究生, 主要研究方向为图像处理、信号处理和深度学习。

E-mail:1098246384@qq.com

闫坤(通信作者), 博士, 教授, 主要研究方向为自适应信号处理、机器学习和图像处理等。

E-mail:kyan5702@gmail.com

郝航勃, 硕士研究生, 主要研究方向为图像处理、信号处理。

E-mail:1628759348@qq.com