

多功能智能导向定位拐杖的研究

林春燕, 卢明杰, 李健松, 陈益炜, 陈源, 刘晶*

五邑大学数学与计算科学学院, 广东 江门

收稿日期: 2026年4月26日; 录用日期: 2026年5月19日; 发布日期: 2026年5月29日

摘要

现如今, 老年人口与视障人群数量逐年增加。他们的出行安全与自主活动能力早已成为重要的社会民生问题。传统拐杖、盲杖一般仅仅只是基础的支撑作用, 缺乏面对复杂环境或突发危险的处理。在此背景下, 我们融合多种技术, 响应国家号召, 从面对复杂场景下的多维度障碍物精准感知、产品的适老化与人机交互、室内外高精度导航定位、产品长续航要求与硬件系统设计、降低成本与产业落地等五大维度出发, 依次采用YOLOv7-tiny模型、NLG技术、DeepSORT算法、设计供电系统总体架构等方案逐一解决, 研发出具有环境感知、精准定位、智能导向、安全预警功能的智能拐杖, 为解决特殊群体出行等难题做出贡献。

关键词

YOLOv7-Tiny模型, NLG技术, DeepSORT算法

Study on Multifunctional Intelligent Navigation and Positioning Crutch

Chunyan Lin, Mingjie Lu, Jiansong Li, Yiwei Chen, Yuan Chen, Jing Liu*

School of Mathematics and Computational Science, Wuyi University, Jiangmen Guangdong

Received: April 26, 2026; accepted: May 19, 2026; published: May 29, 2026

Abstract

The number of the elderly and visually impaired populations is increasing year by year. Their travel safety and independent mobility have long become important social issues. Traditional walking sticks and blind sticks generally only provide basic support, lacking the ability to handle complex environments or sudden dangers. In this context, integrating multiple technologies and responding to national calls, we have developed an intelligent crutch with environmental perception, precise

*通讯作者。

positioning, intelligent guidance, and safety warning functions. It addresses five key dimensions: multi-dimensional accurate perception of obstacles in complex scenarios, age-friendly product design and human-computer interaction, high-precision indoor and outdoor navigation and positioning, hardware system design with long battery life requirements, and cost reduction for industrial implementation. Specifically, we adopt solutions such as YOLOv7-tiny model, NLG technology, DeepSORT algorithm, and design the overall power supply system architecture to solve these problems one by one, contributing to addressing the travel difficulties of special groups.

Keywords

YOLOv7-Tiny Model, NLG Technology, DeepSORT Algorithm

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

现如今,我国人口老龄化趋势明显[1],老年人口数量逐年剧增,随着科技的发展,电子产品的兴起,视障人群数量也随之增多。但老年群体往往因身体机能衰退,普遍存在着平衡能力下降、反应迟缓等问题,同样视障人群因受视力障碍限制,加之城市盲道被占用情况,使得该类特殊人群日常出行成为困难,他们的出行安全与自主活动能力早已成为重要的社会民生问题。传统辅助器具一般仅仅只是基础的支撑作用,缺乏复杂环境或突发危险的应对能力。研发具有环境感知、精准定位、智能导向、安全预警功能的智能拐杖,为解决弥补传统辅助器具短板、解决特殊群体出行等难题做出贡献。本文的研究具有理论意义如下:通过结合前沿技术,以围绕着复杂场景下障碍物精准识别、低算力设备模型轻量化、多模态人机交互等关键问题为核心,开展创新探索,研究成功不仅在原本领域取得成功,也可为智能机器人、自动驾驶、无障碍导航等相关领域提供技术参考与经验借鉴。本文的研究具有较大的实践意义:一是可提升特殊群体出行安全,二是增强特殊群体的出行自主性与生活质量,三是可减轻相关家庭与社会照料负担,四是推动相关产业升级。本研究有利于彰显出科技人文关怀,和加快和谐友好的无障碍社会的建设。

2. 调查与分析

2.1. 国内外研究现状调查分析

1) 国内研究现状

近年来,我国的智能导向拐杖与相关技术发展迅速,相关研究往往紧密结合老龄化社会需求与国家政策,形成了产学研协同推进的局面[2],高校与企业都围绕着核心功能开展,产品集成了超声波、红外、摄像头等多种设备。如在2024年中,广西民族师范学院由杨子靖讲师带队研发的智能导盲拐杖产品中,创新性融合了有超声波测距系统、MaixII-Dock摄像头等设备[3],在2025年中,黄河交通学院与河南省智能制造技术与装备工程技术研究中心联合研发的智能拐杖控制系统中,集成了超声波测距、红外避障等多种设备[4],都为本项目提供了巨大的参考价值,在技术研究上,他们重点关注轻量化模型的目标检测与多维度数据的融合,同时也利用低功耗硬件保障产品的续航和开发出无障碍人机交互,其产品往往更是适合国内场景的使用。

2) 国外研究现状

通过查阅相关文献，国外在相关领域的进度较早，他们以深度学习为基础，围绕 SLAM，通过成熟的技术，早已形成了完善体系。在相关方向，尤其是欧美、日本等发达国家，他们率先将 obstacle detection、GPS 定位、惯性导航等创新型技术加入于产品中[5]，并通过了医疗认证进入商业化应用，同时高度重视标准化与适老化设计，形成相关安全与性能标准。如，在 2021 年美国斯坦福大学智能系统实验室 Patrick Slade 团队研发的主动式智能导盲杖，创新性融合了二维激光雷达、单目相机、惯性测量单元与 GPS 模块，基于激光 SLAM 技术实现了室内外同时定位与地图构建，可实时输出用户位置、行进方向与周围障碍物信息，相关成果发表于《Science Robotics》期刊[6]，在 2022 年，日本的研究团队也推出了搭载 AI 的智能导盲杖产品，通过摄像头与轻量化 AI 模型实现斑马线、铁轨等关键交通标识的识别，为视障者的过街安全提供了低功耗、高可靠性的解决方案[7]。这些研究为智能导盲杖的技术路径提供了重要参考，但其往往成本太高，无法适应国内的道路环境适配。

2.2. 问题分析

本研究围绕着多功能智能导向定位拐杖的设计、研发与实现展开，以老年群体与视障人群安全、自主、便捷出行为核心目标，不断开展系统设计、算法优化、硬件集成与功能验证工作。本项目拟解决关键问题主要分为 5 大点：第一面对复杂场景下的多维度障碍物精准感知问题，第二是产品的适老化与人机交互问题，第三室内外高精度导航定位问题，第四是产品长续航要求与硬件系统设计，第五是降低成本与产业落地。主要内容见图 1。

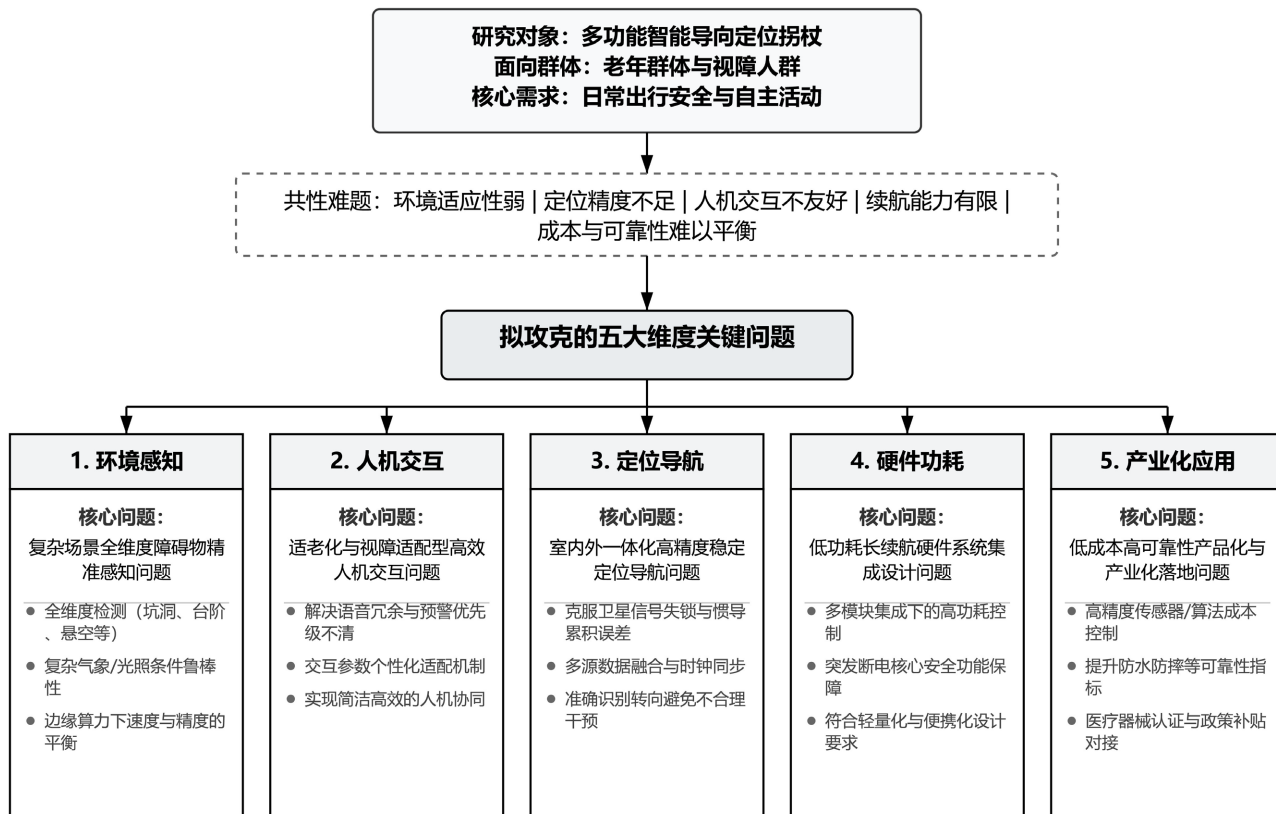


Figure 1. Diagram of key problems to be solved

图 1. 拟解决关键问题图

3. 核心算法设计

3.1. 针对复杂场景下的多维度障碍物精准感知

在面对拐杖在复杂环境下难以避障问题下，我们决定采用视觉与超声波测距相融合的感知设计方案解决该问题，问题解决设计见图 2。

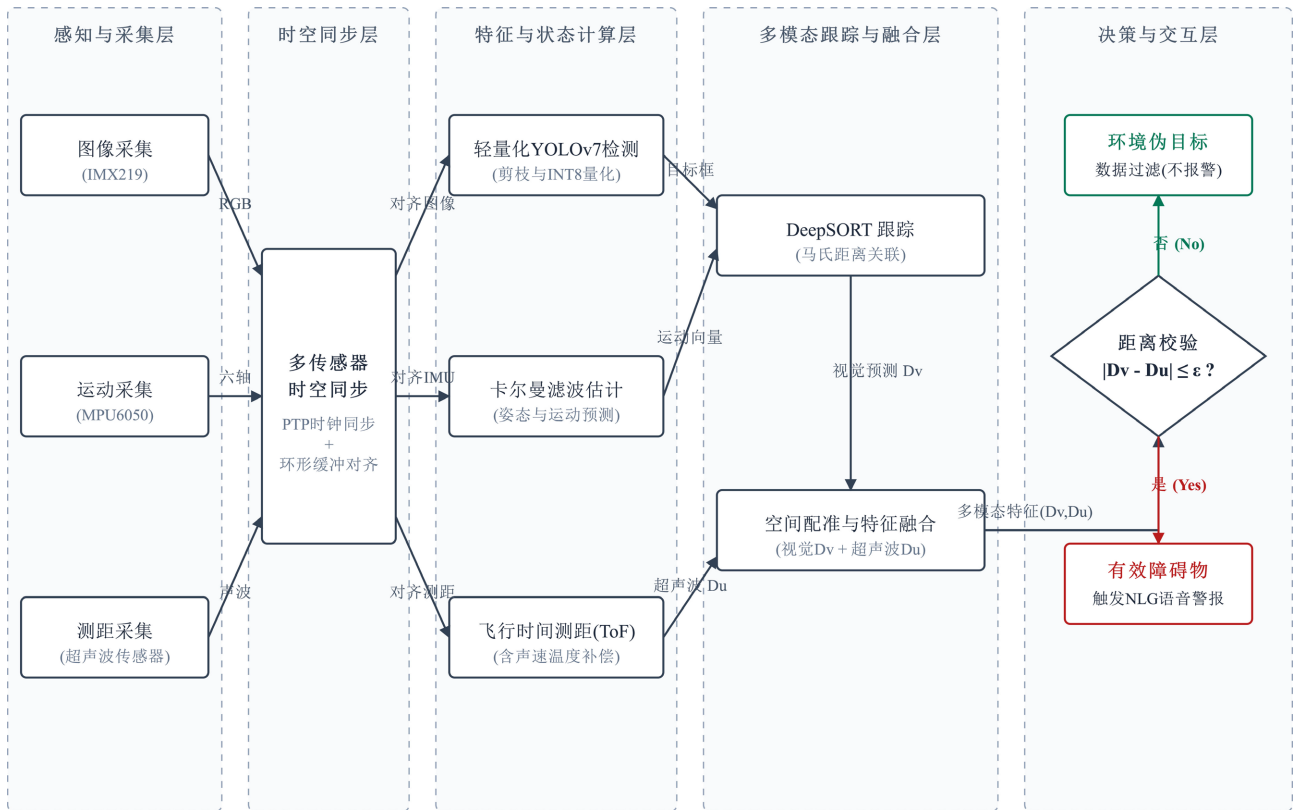


Figure 2. Flowchart of obstacle detection and Multi-modal perception system

图 2. 障碍物检测与多模态感知系统流程图

为此以 Jetson Nano 4GB 为主控核心，构建出“视觉 + 测距 + 惯导”的三维感知架构，见表 1 所示。

Table 1. 3D perception architecture of “Vision + Ranging + Inertial Navigation”

表 1. “视觉 + 测距 + 惯导”的三维感知架构

模块	硬件配置	模块作用
视觉感知模块	IMX219 800 万像素摄像头 (170°广角, 1080P@30fps)	用于获取前方环境图像，基于轻量化 YOLOv7 算法识别行人、车辆、台阶、坑洞等危险目标。
超声波测距模块	超声波传感器	针对 2 cm~4 m 范围内的近距离障碍物，作为视觉结果的二次验证，有效降低复杂环境下的误检率。
运动感知模块	MPU6050 六轴传感器 (加速度计 + 陀螺仪)	采集用户步态及拐杖姿态数据，实时辅助目标轨迹预测。

其次对 YOLOv7 模型进行剪枝与量化优化为 YOLOv7-tiny，利用对 YOLOv7 中冗余的卷积层通道进行 30%的剪枝操作，并利用 TensorRT 进行 FP32 到 INT8 的低精度量化[8]。对模型的压缩与量化加速，

其量化的数学映射关系为:

$$X_{int8} = \text{clip} \left(\text{round} \left(\frac{X_{fp32}}{S} \right) + Z, -128, 127 \right),$$

其中 S 为缩放因子(Scale), Z 为零点偏移量(Zero-point)。

为了稳定追踪目标, 避免因目标遮挡导致的检测丢失, 系统将引入 DeepSORT 算法[9], 并结合 MPU6050 数据进行卡尔曼滤波运动状态预测, 卡尔曼滤波状态更新方程如式(1), 符号说明见表 2。

$$\begin{aligned} X_{k|k-1} &= FX_{k-1|k-1} + BU_k, \\ P_{k|k-1} &= FP_{k-1|k-1}F^T + Q. \end{aligned} \quad (1)$$

Table 2. Symbol description of Kalman filter state update equation

表 2. 卡尔曼滤波状态更新方程符号说明

符号	变量说明
F	状态转移矩阵
B	控制输入矩阵
P	误差协方差矩阵
Q	系统过程噪声协方差

采用马氏距离对卡尔曼滤波的预测轨迹与当前帧 YOLOv7-tiny 的检测结果进行空间匹配度量如式(2), 符号说明见表 3。

$$d^{(1)}(i, j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i). \quad (2)$$

Table 3. Symbol description of Mahalanobis distance spatial matching metric equation

表 3. 马氏距离的空间匹配度量方程符号说明

符号	变量说明
d_j	第 j 个检测框的观测向量
y_i	第 i 个跟踪器在观测空间的投影
S_i	观测空间的协方差矩阵

为提升复杂场景下障碍物感知的稳定性与实时性, 本文在感知层采用“视觉 + 超声波 + IMU”的多传感器融合方案。

在时间同步上, 系统以统一时基进行多源采样管理, 通过 PTP 时间同步机制为图像帧、超声波测距数据与 IMU 数据统一时间戳[10], 并利用环形缓冲区完成异步数据在融合时刻的对齐。具体而言: 视觉数据提供障碍物类别与粗距离信息, 而超声波传感器进行对近距离目标进行二次测距校验, IMU 输出拐杖姿态与运动向量用于动态补偿与轨迹连续性约束。

在空间融合上, 系统先完成相机与超声波到机体系的外参标定, 将视觉距离估计 D_v 与超声波测距 D_u 映射到统一参考系后进行一致性判断。针对视觉检测与轨迹预测的关联问题, 采用基于马氏距离的门控匹配策略。该策略可在遮挡、密集目标与短时漏检场景下抑制误关联, 提高目标级融合鲁棒性。

最终, 系统在“时间对齐 - 坐标配准 - 距离校验 - 门控关联”的闭环流程下输出有效障碍物目标, 为后续语音预警与导航决策提供可靠输入。

3.2. 针对产品的适老化与人机交互

在面对产品的适老化与人机交互问题上，我们决定引入自然语言生成 NLG 技术与文本到语音 TTS 技术[11]，构建“场景结构化提取 - 多目标优先级决策 - 自然语言生成 - 低延迟语音播报”的人机交互架构，见图 3。

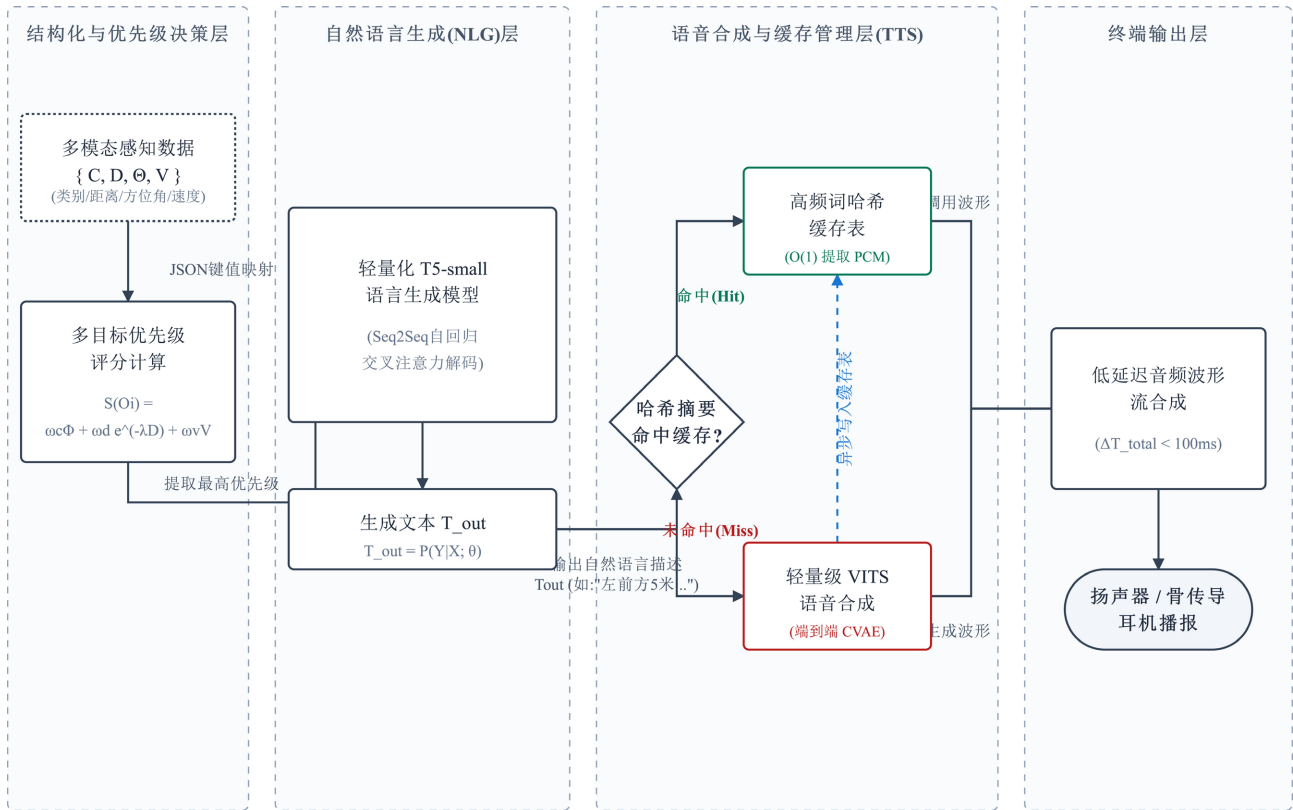


Figure 3. Flowchart of scene understanding and voice prompt system based on NLG
图 3. 基于 NLG 的场景理解与语音提示系统流程图

首先我们通过构建优先级决策模型，对场景信息结构化和确保紧急危险优先播报。建立目标优先级评分模型式(3)，符号说明见表 4。

$$S(O_i) = \omega_c \cdot \Phi(C_i) + \omega_d \cdot \exp(-\lambda D_i) + \omega_v \cdot \max(0, V_i \cos \Theta_i), \quad (3)$$

Table 4. Symbol description of target priority scoring model
表 4. 目标优先级评分模型符号说明

符号	变量说明
$\Phi(C_i)$	类别的基础危险系数(例如，车辆、坑洞的 Φ 值赋为最高，行人次之)
$\exp(-\lambda D_i)$	距离衰减函数，距离越近威胁值呈指数上升(λ 为距离衰减常数)
$V_i \cos \Theta_i$	目标朝向用户的径向接近速度
$\omega_c, \omega_d, \omega_v$	归一化权重参数 $\omega_c + \omega_d + \omega_v = 1$

提取 $S(O_i)$ 最高的目标作为主导障碍物输入到语言生成模块，从而实现对紧急障碍物的优先响应。

其次接入轻量化 T5-small 自然语言生成模型，将其部署于边缘计算平台进行场景语言描述的动态生成。采用低延迟语音合成与多级缓存机制，实现基于 VITS 的轻量化语音合成，给定文本表示 c ，其目标是最大化音频波形 x 的条件对数似然下界(ELBO)为：

$$\log p_{\theta}(x|c) \geq \mathbb{E}_{q_{\phi}(z|x)} [\log p_{\theta}(x|z)] - D_{KL}(q_{\phi}(z|x) \| p_{\theta}(z|c)).$$

采用高频提示语句缓存策略，在 TTS 前端构建了 $O(1)$ 时间复杂度的哈希缓存表(Hash Cache)。令当前 NLG 输出文本为 T_{out} ，系统首先计算其哈希摘要为：

$$H_{key} = MD5(T_{out}).$$

实现人机交互，同时确保人机交互的响应速度与系统续航能力。

3.3. 针对解决室内外高精度导航定位问题

针对该问题，我们决定采用基于北斗与微机电惯性测量单元的松耦合组合导航架构，随后通过引入用户意图识别模型，实现用户主动转向与路径偏离精准区分，见图 4。

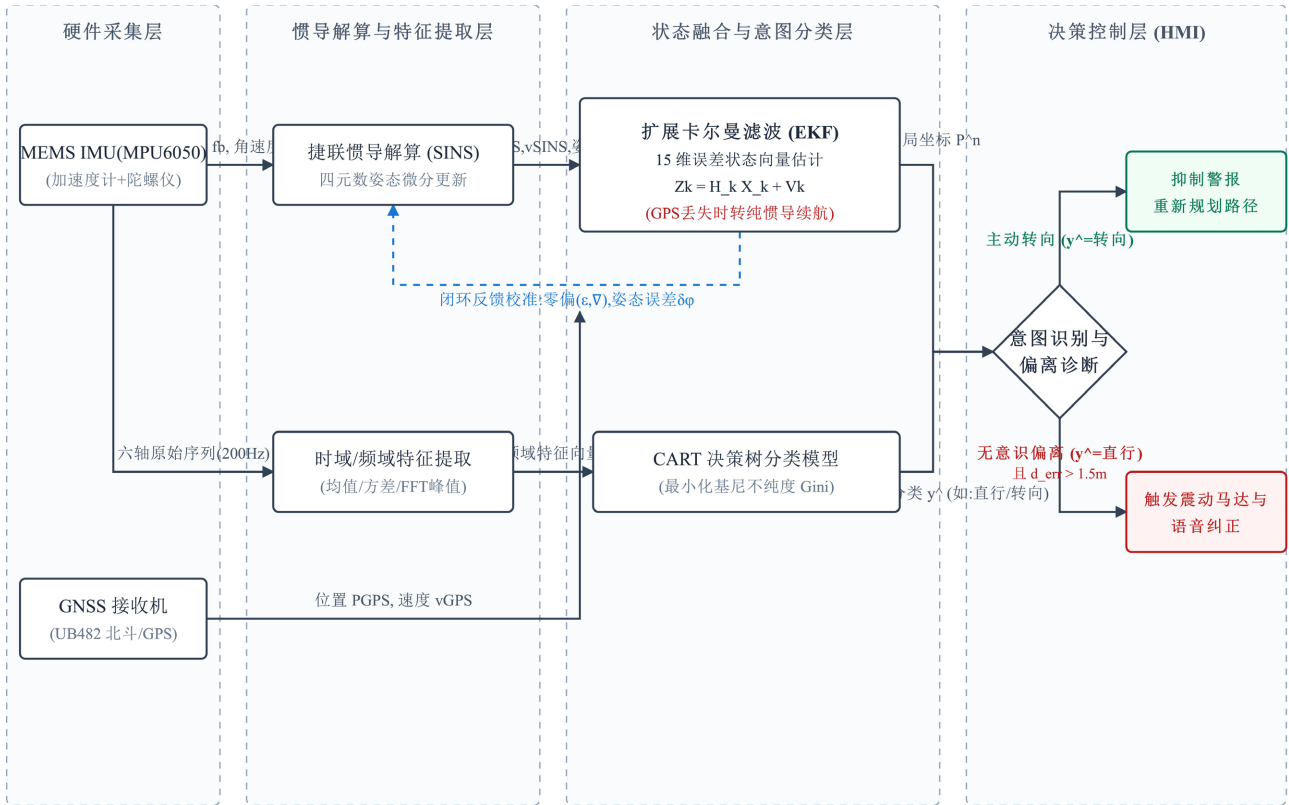


Figure 4. Flowchart of navigation assistance system based on BDS/IMU fusion and intent recognition
图 4. 基于北斗导航系统/IMU 融合与意图识别的导航辅助系统流程图

首先我们选用芯星通 UB482 高精度北斗三号与 MPU6050 六轴传感器作为硬件支持，以极高的采样率(50 Hz)捕获拐杖的瞬时运动学参数。其次构建扩展卡尔曼滤波模型进行航位推算优化[12]。

定义系统的 15 维误差状态向量 X 为：

$$X = [\delta\phi^n, \delta v^n, \delta P^n, \epsilon^b, \nabla^b]^T,$$

其中 $\delta\phi^n$ 为姿态误差、 δv^n 为速度误差、 δP^n 为位置误差、 e^b 为陀螺仪随机游走漂移与 ∇^b 为加速度计零偏。离散化后的系统状态方程为：

$$X_k = \Phi_{k,k-1} X_{k-1} + W_{k-1},$$

其中， $\Phi_{k,k-1}$ 为由 SINS 误差微分方程推导得到的一步状态转移矩阵， W_{k-1} 为系统过程噪声序列。

当北斗信号有效时，取北斗输出的位置 P_{beidou} 和速度 v_{beidou} 与 IMU 解算的位置 P_{SINS} 和速度 v_{SINS} 之差作为 EKF 的观测向量 Z_k ：

$$Z_k = \begin{bmatrix} P_{SINS} - P_{beidou} \\ v_{SINS} - v_{beidou} \end{bmatrix} = H_k X_k + V_k,$$

其中 H_k 为测量矩阵， V_k 为 GPS 观测噪声。

噪声协方差矩阵的设定遵循“传感器物理噪声 + 场景自适应”原则[13]：过程中噪声协方差 Q 依据 IMU 的角速度噪声、加速度噪声及零偏随机游走特性进行离散化构造；观测噪声协方差 R 依据北斗定位质量动态调整，在开阔环境取较小值，在遮挡或多路径环境适当放大。初始协方差 P_0 则根据开机对准误差、初始速度不确定度与初始定位精度设定。该建模方式可在 GNSS 短时不稳定或中断时保持惯导连续性，并在信号恢复后快速收敛，实现室内外过渡场景下的稳定导向定位。

最终基于 500 小时的盲人真实行走测试数据集，构建了包含 12 种步态模式(如平地直行、原地转向、侧步避让、上下台阶等)的 CART (Classification and Regression Tree) 决策树模型对用户意图与步态进行识别。

3.4. 针对低功耗硬件系统集成与结构设计

在智能助行辅具的工程设计中，能源系统的可持续性与稳定性是决定产品实用性与安全性的核心指标。本产品主要面向老年与视障人群全天候户外出行的使用场景，其集成了视觉感知、多传感器检测、北斗定位等多类功能模块，对能源系统提出了长续航、低功耗、高稳定、强安全、易操作的综合设计要求[14]。

首先我们对整机功耗构成进行区分，区分为静态功耗与动态功耗。明确各模块能耗分布规律，为低功耗设计提供量化依据：静态功耗源于电路漏电流与模块待机能耗，在设备休眠、静止等待状态下持续存在，主要由传感器待机、主控休眠电路产生；动态功耗与系统运行频率、算力负载、模块工作状态直接相关，是整机能耗的核心组成部分。

Table 5. Analysis table of overall power consumption composition and characteristics of the intelligent navigation and positioning crutch

表 5. 智能导向定位拐杖整机功耗构成与特性分析表

模块分类	核心硬件	功耗特征分布	运行模式与能耗表现	优化潜力
主控计算模块	Jetson Nano	系统最大能耗源；算力与功耗呈正相关	全速模式：满载功耗达峰值 待机模式：能耗大幅下降	极高(算力调度、模型轻量化)
视觉感知模块	IMX219 摄像头	持续工作外设；能耗输出稳定，占比固定	全速模式：高频采样 待机模式：关闭高频运算	中等(启停控制、采样率调节)
基础传感模块	超声波 MPU6050	静态电流微安级；整体能耗占比极低	常规模式：全时段运行 待机模式：硬件休眠，快速唤醒	较低(选用低功耗器件已达标)
扩展交互模块	北斗定位 语音振动	间歇式工作机制；常态待机能耗可忽略	预警模式：按需瞬间启动 常规模式：待机断电	高(供电切断、降低采样频率)

其次进行特性分析，从主控算力调度、视觉模块启停、多模块间歇工作作为核心切入点，在不削弱安全功能的前提下，显著降低整机平均功耗，提升能源利用效率，见表 5。

然后针对整机功耗分布特征，采用硬件分级供电 + 软件动态调度的协同优化方案，从硬件架构、器件选型、算法调度出发，进行多维度降能耗，实现性能与续航的平衡，见表 6。

Table 6. Low-Power hardware and software collaborative optimization strategies
表 6. 低功耗硬件协同优化策略表

优化维度	核心策略	具体技术实施方案	预期技术效果
硬件优化	分级供电架构	划分核心安全模块(常供电)与扩展模块(可控供电)，非必要场景切断扩展电源	杜绝无效能耗，延长单次续航
硬件优化	器件与电路优化	选用工业级微安级待机器件；采用低阻抗布线与高效 DC-DC/LDO 稳压转换	降低静态基线，提升电能转换效率
软件优化	动态状态感知与调频	基于 MPU6050 步态数据识别状态，静止时自动降频并关闭视觉采样，行走满载	实现整机能耗随应用场景自适应调节
软件优化	模块间歇调度	降低非预警时段超声波测距与北斗定位采样频率	减少外设持续工作能耗，保证检测精度
算法优化	模型算力轻量化	采用 YOLOv7-tiny 模型剪枝、INT8 量化与 TensorRT 加速技术	削减主控推理峰值功耗，兼顾实时性

在结合设备功耗特性与户外使用需求，构建可拆卸锂电池主供电、超级电容应急供电、PMU 智能管理的一体化供电架构，实现长续航、高安全、稳输出的能源供给[15]，见图 5。

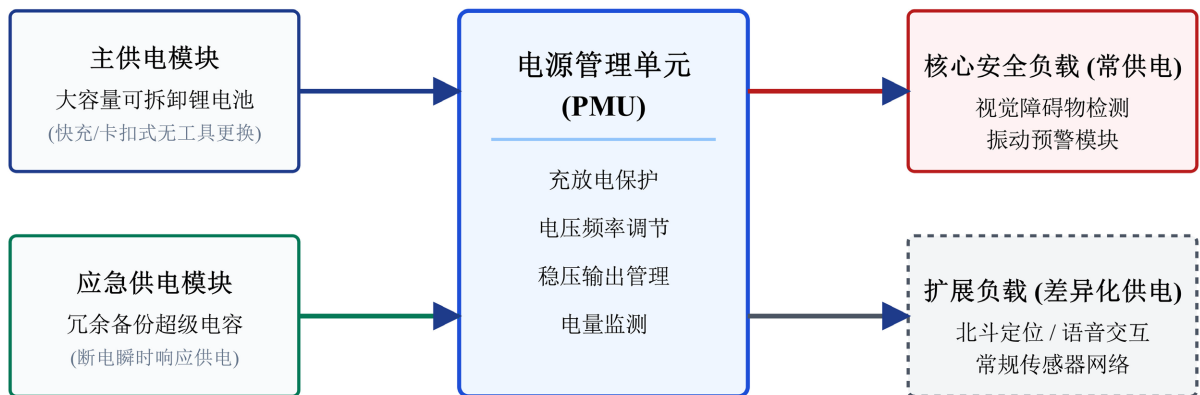


Figure 5. Overall topology architecture of the power supply system
图 5. 供电系统总体拓扑架构

最后，可靠性与适老化是助行辅具能源系统的核心设计准则，我们从环境防护、安全保护、老化适配、操作优化四个维度，全面提升了系统全生命周期稳定性与易用性，见表 7。

Table 7. Decomposition table of reliability and age-friendly design indicators for the energy system
表 7. 能源系统可靠性与适老化设计指标分解表

设计维度	二级指标	技术实现手段与规范	解决的用户痛点
可靠化设计	环境耐受防护	供电模块支持宽温工作； 电池与电路做防水、防摔、抗电磁干扰封装	户外高低温、潮湿等 复杂环境适应性差

续表

可靠化设计	电气安全矩阵	PMU 集成过压、过流、短路、过温四重保护；严控充放电阈值	避免过充过放导致的硬件损毁与安全隐患
可靠化设计	故障监测机制	实时监测电压电流与电池健康状态，异常触发语音预警	提升设备可维护性与突发故障安全性
适老化设计	极简交互体验	采用防盲插充电接口，全程无按键，NLG 生成语音播报剩余电量	视障与高龄人群精细操作困难、视觉读取困难
适老化设计	电池老化补偿	软件算法实时评估健康度，动态调整策略补偿容量衰减	长期使用后电池内阻上升导致的续航缩短
适老化设计	结构轻量易护	电池无工具卡扣式快速更换；供电模块高集成度轻量化	降低维护门槛，保障握持舒适性，不增负重

3.5. 降低成本与产业落地

紧跟国家相关政策[16]，让产品紧跟潮流，同时开展第三方企业合作路线的探索[17]，实现产品的量化生成与自己的产业链，助力无障碍社会的建设。

4. 实验测试

4.1. 实验过程

1) 数据的划分与参数的设计

在实验的训练数据方面上，首先我们通过选取三类典型公开场景：Argoverse、VisDrone 与 xView 作为实验数据，将其进行统一转换为 YOLO 标注格式，并完成训练集、验证集划分。其次对数据类别名称、标签完整性和异常样本进行人工核查，以确保训练数据质量一致。

在实验参数设计上，我们通过结合相关资料，首先将固定随机种子设置成降低随机扰动，对模型输入的尺寸设为 640×640 [18]，训练周期设为 300 *epoch*，*batch size* 设为 6，优化器采用 SGD [19]，其余参数则保持与基线一致。其次在评价指标上，我们采用 *Precision (P)*、*Recall (R)*、*mAP@0.5* 与 *mAP@0.5:0.95*，其中以 *mAP@0.5:0.95* 作为主指标 [20]，并从三方面进行评估：语义维(多类别障碍物检测)、几何维(旋转框感知能力)与时序维(检测 - 跟踪连续感知能力)

2) 基线模型训练与本文方法训练

在此我们先训练基线检测模型 YOLO11n，通过记录每个 *epoch* 的 *box_loss*、*cls_loss*、*dfl_loss* 与验证集 *P*、*R*、*mAP@0.5*、*mAP@0.5:0.95*，并保存最优权重用于后续对比，见表 8 所示。

Table 8. Optimal weights during the experiment

表 8. 实验过程中最优权重

指标	数值
<i>best_epoch</i>	276
<i>best_P</i>	0.864
<i>best_R</i>	0.780
<i>best_mAP@0.5</i>	0.857
<i>best_mAP@0.5:0.95</i>	0.599

我们在基线基础上，逐步引入多尺度训练、复杂场景增强、几何维建模(OBB)与时序一致性优化，对

模型采用与基线相同的数据划分和训练策略，以保证对比公平。

3) 主实验评估与消融实验设计

针对主实验评估，我们在三个数据集与验证集上，分别评估基线、对比方法与本文方法，统一采用评价指标 P 、 R 、 $mAP@0.5$ 和 $mAP@0.5:0.95$ ，进行统计，并以跨数据集平均性能作为综合指标。

针对消融实验，为了验证方法关键环节的有效性，在此我们通过消融实验，逐步去除或加入各个模块，明确性能提升的真实来源，从而避免将结果误判为训练偶然性或参数巧合[21]。我们以“Baseline→+ 多尺度→+ 增强策略→+ OBB→+ 时序优化”的渐进方式开展消融，逐项量化各模块对性能提升的贡献，验证模块必要性与协同增益。

4) 时序维评估与可视化与误差分析

在真实复杂场景下，障碍物通常会因为未知因素而发生遮挡、出入画、尺度变化和快速运动。为此我们需要通过时序维评估，以 MOTA、IDF1、IDs 为时序指标[22]重点分析遮挡、快速运动与目标出入画等动态场景下的连续感知稳定性。同时对典型场景输出进行检测其可视化、PR 曲线与混淆矩阵，统计误检与漏检样例，重点分析密集小目标、强遮挡和长尾类别中的失败模式，并总结方法边界与后续改进方向。

4.2. 实验结果

根据上述实验步骤，我们对关键实验至少重复 3 次及以上，并取其平均值，最终形成主实验表(见表 9)、消融表(见表 10)与时序评估表(见表 11)。最终给出相关的结论如下。

Table 9. Record table of main experiment results

表 9. 主实验结果记录表

方法	Argoverse mAP50-95	VisDrone mAP50-95	xView mAP50-95	平均 mAP50-95
YOLOv8n	0.621	0.438	0.392	0.484
YOLO11n	0.647	0.467	0.421	0.512
RT-DETR-R18	0.659	0.451	0.433	0.514
本文方法	0.681	0.512	0.471	0.555

将根据实验指标，对实验结果进行统计，分析得出本文方法与次优方法相比，在三类复杂场景均取得更高主指标，说明该方法对遮挡、密集小目标与长尾类别具有更好的鲁棒性，同时在 VisDrone 与 xView 上提升更明显，在高难度场景中优势更加突出明显。

Table 10. Record table of model ablation experiment results

表 10. 模型消融实验结果记录表

配置	Argoverse	VisDrone	xView	平均
Baseline (YOLO11n)	0.647	0.467	0.421	0.512
+多尺度训练	0.661	0.486	0.439	0.529
+复杂场景增强策略	0.672	0.501	0.456	0.543
+几何维建模(OBB 分支)	0.676	0.506	0.463	0.548
+时序一致性优化(最终)	0.681	0.512	0.471	0.555

基于该实验结果，我们分析可得多尺度训练与复杂增强策略是模型性能提升的主要来源[23]，且在几

何维建模中我们进一步降低了斜向目标和密集目标的定位误差。而时序一致性优化在视频流场景中可以减少瞬时误检，使最终性能达到最优。

Table 11. Record table of model temporal dimension evaluation results
表 11. 模型时序维评估结果记录表

方法	MOTA	IDF1	IDs	FPS
ByteTrack + Baseline	67.4	69.1	412	56.3
BoT-SORT + Baseline	68.2	70.4	389	49.7
BoT-SORT + Proposed	71.6	74.2	301	47.9

可知模型在复杂动态场景下表现出更强的连续感知能力。

实验结果表明，本文方法在车路、低空密集目标与遥感长尾等复杂场景中均实现稳定增益，且通过分阶段消融，验证了多尺度训练、复杂增强、几何维建模与时序一致性优化对性能提升的贡献。说明其性能提升来源清晰、复杂场景下感知能力有效且具备实际部署价值。

5. 实验结果分析

本实验在 Argoverse、VisDrone 与 xView 三类数据集上进行对比实验，通过统一采用 $mAP@0.5:0.95$ 作为主指标。结果显示，本文方法在三类场景上均取得最优性能：Argoverse 为 0.681、VisDrone 为 0.512、xView 为 0.471，跨数据集平均值达到 0.555。相较 YOLO11n 基线的 0.512，提升为+0.043，相较次优方法的 0.514 提升为+0.041。

在消融实验中，我们采用“Baseline→ + 多尺度训练→ + 复杂场景增强→ + OBB 几何建模→ + 时序一致性优化”的渐进方式，性能由 0.512 稳定提升至 0.555，验证了各模块的有效性与协同作用。

时序评估中，BoT-SORT + Proposed 的 MOTA 为 71.6、IDF1 为 74.2、IDs 为 301，优于 BoT-SORT + Baseline (MOTA68.2, IDF170.4, IDs389)，说明模型在动态场景中具有更强的连续感知能力[24]。全部关键实验均重复至少 3 次并取平均，结果具有一定稳定性。

从结果可见，多尺度训练与复杂增强策略是性能提升的主要来源，OBB 几何建模进一步改善了斜向目标和密集目标定位。时序一致性优化有效降低了视频流中的瞬时误检与身份切换。本文方法在遮挡、密集目标与长尾类别场景中的优势更为明显，体现出较好的复杂环境鲁棒性与跨场景泛化能力。同时，FPS 从 49.7 升至 47.9，表明系统以小幅实时性代价换取了更高时序稳定性，该权衡符合智能助行场景对安全与连续性的需求。后续我们将围绕全场景长周期验证、标准化与安全认证、以及低功耗与实时性协同优化持续推进，以提升系统的工程可落地性与实际应用价值。

6. 结论

本项目主要以老年人群体与视障人群需求为出发点，以满足其安全、自主、便捷出行为核心目标展开，通过围绕多功能智能导向定位拐杖展开研究，目前完成工作主要有系统的设计与开发，视觉算法的研发，相关硬件的集成与设计方案验证。本项目通过构建一套能够多维度识别全环境、室内外精确定位、便捷性的人机交互系统、长时间低功耗的续航的体系，有效的弥补了市场上传统拐杖的缺点，使得产品更加贴合我国真实情况。本研究虽完成了既定的核心研发与验证任务，但结合工程落地与全场景适配要求，仍存在缺乏大规模、长周期实地测试未完成医疗器械认证，未对接助残养老补贴政策等不足，未来，在面向无障碍社会建设与智慧养老产业的发展趋势中，本团队将从技术优化、功能拓展、硬件升级、产业化落地四个维度，不断对项目进行持续性完善，为推动智能拐杖向更实用、更普惠、更可靠的方向发展。

基金项目

五邑大学 2025 年校级大学生创新创业训练计划项目“多功能智能导向定位拐杖”(X202511349012)。

参考文献

- [1] 国家统计局. 第七次全国人口普查公报[R]. 北京: 国家统计局, 2021.
https://www.stats.gov.cn/sj/tjgb/rkpcgb/qgrkpcgb/202302/t20230206_1902008.html, 2026-04-05.
- [2] 刘静, 赵宇. 轻量化深度学习在视障碍检测中的应用研究综述[J]. 计算机应用研究, 2024, 41(3): 641-649.
- [3] 黄玉瑶, 杨子靖, 陈慧萱, 等. 基于摄像头的智能导盲拐杖[J]. 电子制作, 2024, 32(22): 49-52.
- [4] 杨蒙蒙, 张兆宽, 董浩浩, 等. 智能拐杖控制系统设计[J]. 电子制作, 2025, 33(18): 27-32.
- [5] Redmon, J. and Farhadi, A. (2023) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, 18-22 June 2023, 1-10.
- [6] Slade, P., et al. (2021) Improved Mobility for People with Visual Impairments Using an Active Smart Cane with Simultaneous Localization and Mapping. *Science Robotics*, 6, eabj5066.
- [7] 日本情報処理学会. スマート盲導杖「みちしる兵衛」～AI 搭載白杖による視覚障害者歩行支援～[C]//第 84 回情報処理学会全国大会講演論文集. 2022: 8018.
- [8] 季鑫焯. 基于改进 YOLOv7-tiny 的航拍目标检测与轻量化设计[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2025.
- [9] 礼冬雪, 张宁, 王宏宇, 等. 基于改进的 YOLO 和 DeepSORT 的多目标定位与跟踪算法[J]. 无人系统技术, 2025, 8(6): 122-132.
- [10] 易月娥. 低功耗状态下无线传感器网络的定位和时间同步机制研究[J]. 湖南邮电职业技术学院学报, 2022, 21(3): 21-24.
- [11] 杨文钰. 人机交互范式下自然语言生成技术的社交媒体传播效果量化研究[J]. 互联网周刊, 2025(16): 36-38.
- [12] 孙帅. 误差状态卡尔曼滤波 SINS/GPS 组合导航方法研究[D]: [硕士学位论文]. 西安: 西安理工大学, 2024.
- [13] 刘飞, 姚晖. 基于干扰加噪声协方差矩阵重构的嵌套阵列自适应波束形成算法[J]. 通信技术, 2025, 58(5): 471-478.
- [14] 国家康复辅具研究中心. 中国智能康复辅具产业发展报告(2023) [M]. 北京: 机械工业出版社, 2023.
- [15] 黄平, 林晓. 可穿戴助行设备续航与节能技术研究[J]. 电源技术, 2023, 47(3): 341-344.
- [16] 国务院关于印发“十四五”残疾人保障和发展规划的通知[Z/OL]. 2021.
https://www.gov.cn/zhengce/content/2021-07/21/content_5626391.htm, 2026-04-05.
- [17] 柯园, 吴茂璠, 应雅梦, 等. 基于“互联网+”的智能拐杖创新研究[J]. 现代商贸工业, 2018, 39(35): 68-70.
- [18] 张新航, 张雅茹, 麻振华, 等. 基于深度学习方法的 YOLO 目标检测综述[J]. 长江信息通信, 2024, 37(8): 52-56.
- [19] 秦立峰, 李博株, 林敬轩, 等. 基于 SGD-YOLO 模型的黄瓜霜霉病检测[J]. 农业机械学报, 2026, 57(2): 203-214.
- [20] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2022) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. <https://arxiv.org/abs/2207.02696>
- [21] Zeiler, M.D. and Fergus, R. (2014) Visualizing and Understanding Convolutional Networks. In: Fleet, D., et al., Eds., *European Conference on Computer Vision*, Springer, 818-833. https://doi.org/10.1007/978-3-319-10590-1_53
- [22] 王相威. 复杂场景下的多目标跟踪算法研究[D]: [硕士学位论文]. 杭州: 杭州电子科技大学, 2025.
- [23] 苗丁锐博, 鄢建国, 涂志刚, 等. 窄角相机月球影像自适应多尺度超分辨率重建[J]. 中国图象图形学报, 2025, 30(9): 3050-3065.
- [24] 王尔申, 顾佳怡, 徐嵩, 等. 基于 Deep SORT 的多无人机跟踪优化算法[J]. 电光与控制, 2026, 33(4): 1-6.