

# 基于 RK3588 的智能酒店服务机器人

## 摘要

本项目基于 ELF2 平台，设计并实现了一款面向酒店服务场景的智能机器人“UBoot”。系统融合语音识别、自主导航、目标检测与环境感知等多模态智能技术，旨在满足智慧酒店对引导、服务与安全辅助的高可靠性需求。

机器人使用 **ROS Foxy+ubuntu20.04** 系统部署建图，导航，语音等任务，使用 RK3588 部署视觉检测算法，充分释放算力潜力。

**导航**方面，系统融合激光雷达、IMU 与里程计数据，基于 Cartographer 实现高精度 SLAM，结合 ROS 2-Nav2 构建路径规划、行为树管理与局部避障的完整导航链，适应酒店复杂结构环境。

**语音**方面，系统采用本地 Vosk 模型，Spark 4.0 云服务，REST 接口实现语音识别，语义理解，语音合成，识别准确率达 95%，延迟低于 3 秒，构建出高效、流畅的智能语音交互链路。

**视觉**方面，系统部署轻量化 YOLOv5-Lite 模型，支持瓶子、纸巾、烟雾等多类常见垃圾目标的实时检测，mAP@0.5 达 94%，帧率达 28–30 FPS，具备良好的光照适应性与鲁棒性。

综上所述，本项目在 **RK3588 国产平台**上实现了导航、语音与视觉等核心功能的高效协同，构建出一套面向酒店场景的智能服务机器人系统。系统具有良好的实时性、稳定性与多模态感知能力，展示了国产软硬件在智慧服务领域的应用潜力，具备实际部署与推广价值。

关键词：ROS2，自主导航，语言大模型，多模态，RK3588

## 第一部分 作品概述

### 1.1 功能与特性

本系统面向酒店服务场景，围绕“智能化、自主化、实用性”进行功能模块设计，主要包括以下几个核心功能：

#### (1)语音识别与情感交互

系统集成百度语音识别和讯飞语音合成 API，实现语音唤醒、语义理解与情感反馈。结合情绪识别模型，判断用户情绪，提升交互体验。

#### (2)自主导航与带路服务

融合激光雷达、IMU、里程计，实现高精度定位与建图，基于 ROS 2 导航算法自主带路，支持动态避障与路径重规划，保障安全高效。

#### (3)垃圾检测与环境感知

利用多种传感器实时监测环境，CPU 处理基础数据，NPU 处理视觉任务，实现垃圾识别与环境适应，提升服务智能化水平。

### 1.2 应用领域

在人工智能与智能硬件快速发展的背景下，服务机器人逐渐成为推动智慧酒店转型升级的重要力量。从图 1 中可以看出，服务机器人已广泛应用于接待、引导、清洁等多个酒店服务环节，正加速推动传统酒店向智能化、自动化方向转型。

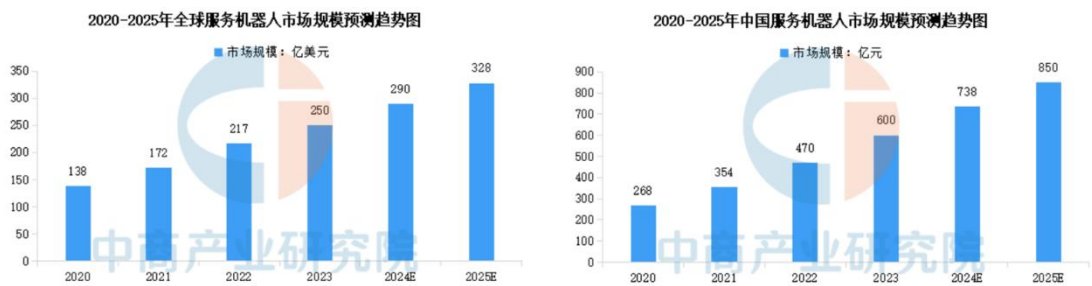


图 1 服务机器人销量图

传统酒店服务机器人普遍存在适用范围局限、环境适应能力差、缺乏多模态感知与交互单一等问题，难以满足复杂场景下的稳定运行与智能服务需求。为解决上述问题，本项目基于国产 ELF2 开发板，设计并实现了一款具备自主导航、语音情感交互和目标识别能力的“UBoot”酒店服务机器人。

### 1.3 主要技术特点

本项目围绕“智能化、多模态、国产适配”三大核心理念，构建面向酒店场景的智能服务机器人系统，具备以下技术特点：

**(1) 异构算力协同调度：**基于 ELF2 平台，系统充分利用高性能 CPU 与 RK3588 NPU 资源，采用“CPU 执行调度控制 + NPU 执行视觉推理”的异构协同机制，实现智能导航与高效视觉检测的任务隔离与并行加速。

**(2) 多模态感知与语义闭环交互：**融合激光雷达、IMU、语音麦克风阵列和摄像头等多源传感器，支持 SLAM、YOLOv5-Lite 视觉识别、语音识别与语义理解等多模态输入，构建语义驱动的闭环行为响应链条。

**(3) 轻量视觉模型与高效部署：**采用 YOLOv5-Lite 并进行 RKNN 量化优化，部署于 NPU 加速推理，检测帧率达 25-30FPS，mAP@0.5 高达 96.9%，兼顾精度与速度，适应边缘平台资源限制。

**(4) 系统级软硬件协同设计：**自主设计电源与转接板，构建模块化硬件结构；软件采用 ROS 2 通信框架，任务调度器统一管理导航、语音与环境感知，整体系统集成度高、稳定性强，具备良好的工程可部署性与扩展能力。

## 1.4 主要性能指标

本系统集成语音、导航、视觉与多核通信，具备高识别率、低延迟与优良导航性能，满足酒店复杂环境下的实时性与稳定性需求，具有实际部署价值。

表 1 系统性能指标

功能模块	性能指标
自主导航	到点成功率 $\geq 95\%$ ，路径规划时间 $< 1000\text{ms}$
定位精度	定位偏差 $\leq 30\text{cm}$ ，轨迹偏差 $\leq 15\text{cm}$
语音交互	识别准确率 $\geq 85\%$ ，响应延迟 $\leq 3$ 秒
目标检测	Precision 89.8%，Recall 93%，检测帧率 25~30 FPS
稳定运行时长	$\geq 2$ 小时连续运行无异常
系统续航	1 小时续航，可换电池

## 1.5 主要创新点

**(1) 基于 RK3588 的异构算力深度协同架构：**项目充分发挥 ELF2 高性能 CPU 与 RK3588 异构计算能力，构建“CPU 任务调度+NPU 推理加速”机制。智能导航

与语音交互模块运行于 CPU 侧，实时视觉检测部署 RKNN 格式 YOLOv5-Lite 模型于 NPU 推理，实现任务隔离、计算加速与资源最优调度。

(2) 多模态智能融合体系：集成 SLAM、语音理解、目标检测等异构感知链路，构建面向服务任务的语义驱动行为响应机制，支持引导、问询与垃圾识别等复合场景。

(3) 视觉感知高效轻量部署：YOLOv5-Lite 模型经过 SFB 与通道重排优化后，转换为 RKNN 模型部署在 NPU 侧，推理帧率达 28–30FPS，检测精度 mAP@0.5 高达 96.9%，兼顾实时性与鲁棒性。

(4) 自研电路与结构平台：定制电源管理与转接板模块，支持多传感器接入与稳压保护；底盘结构优化布线与维护性，具备良好的系统工程集成能力与部署可靠性。

## 1.6 设计流程

本项目采用“需求分析—模块开发—系统集成—场景验证”四阶段流程。团队首先明确酒店服务场景下对语音交互、导航避障、视觉识别等功能的需求，确立基于 ROS 2+RK3588 加速的技术架构。

系统开发中，语音模块融合本地 Vosk 与 Spark 4.0 语义理解；导航基于 Cartographer+Nav2 构建完整路径规划与避障链路；视觉模块将 YOLOv5-Lite 模型转 RKNN 格式，使用 NPU 实现高效推理。各模块通过 ROS 2 通信集成，最终由任务调度器统一调控多模态行为，实现系统稳定运行与智能协同。

## 第二部分 系统组成及功能说明

### 2.1 整体介绍

本系统基于 ROS 2 Foxy 构建，采用“机器人本体+边缘设备+分布式计算+云服务”四层混合架构，形成高效、可扩展的智能协同体系。下图展示了 Uboot 系统的网络架构和模块间关系。通信核心为局域网中的路由器，连接了本体机器人、分布式计算平台与 ROS 接入的边缘智能设备。

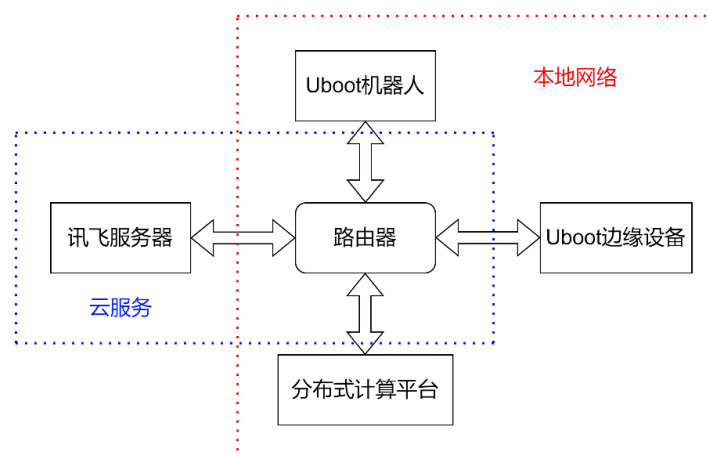


图 2 Uboot 架构图

Uboot 机器人在本地通过 ROS 2 节点完成导航控制、传感器采集与任务执行，充分利用 ELF2 平台的 CPU 的并行能力。

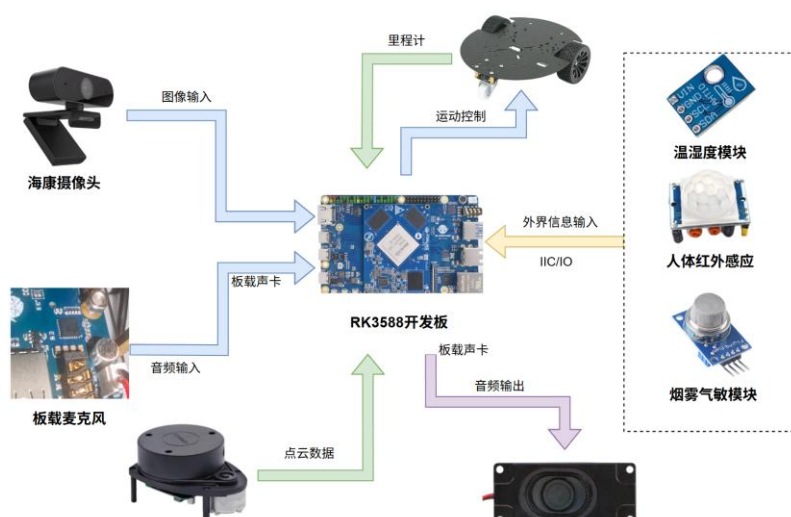


图 3 硬件系统图

系统通过局域网搭建分布式框架，加载语音识别、图像分析等高计算模块，

并连接讯飞云服务，支持情感化语音交互。

## 2.2 硬件系统介绍

### 2.2.1 硬件整体介绍

如图 4 所示，系统基于 ELF2 构建，具备强大的边缘计算能力与丰富接口资源。系统通过 USB 连接海康摄像头、激光雷达与控制板，通过 IIC 和 GPIO 连接温湿度、烟雾传感器等模块。

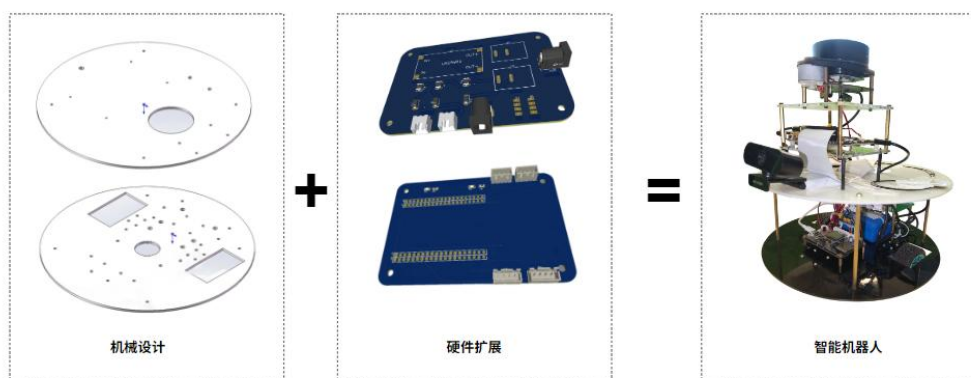


图 4 底盘搭建

本项目采用“机械设计 + 硬件扩展”并行策略，定制机器人结构与传感器安装位，优化散热与维护；同时设计电源板、转接板等核心电路，满足系统集成与稳定运行需求，体现出良好的工程化与集成能力。

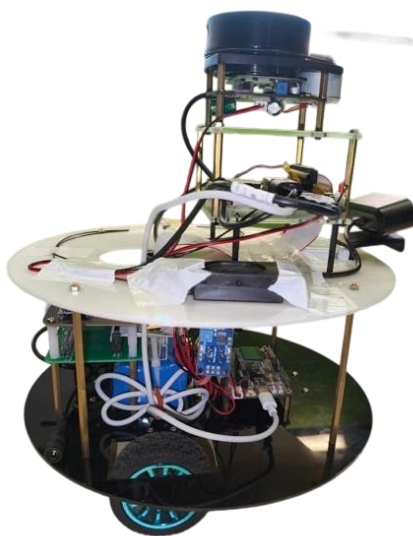


图 5 小车机械结构

### 2.2.2 机械设计介绍



小车底盘直径 280mm，采用两主动轮（左右两侧）+2 个从动万向轮（前后）的差速驱动结构。如图 6 为小车的亚克力层板。

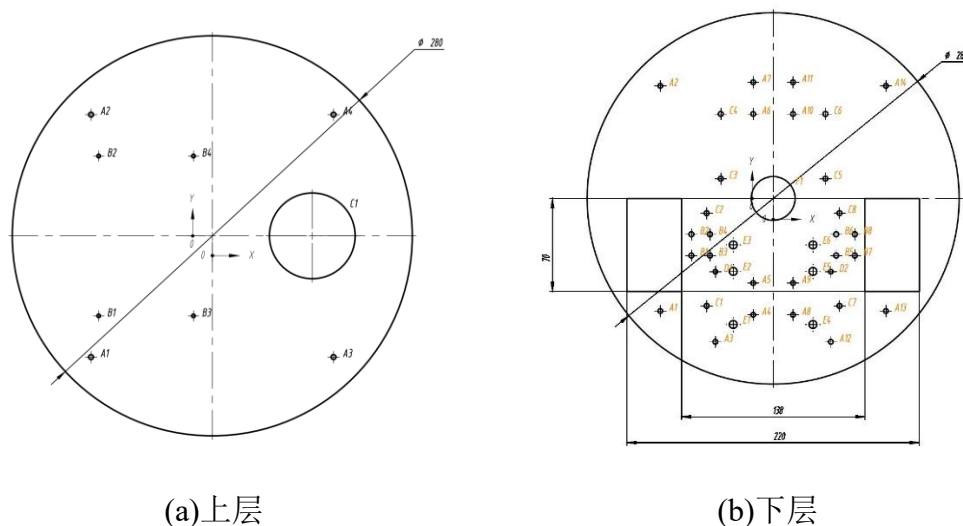


图 6 小车层板

上层平台安装 ELF2 开发板和激光雷达，摄像头，喇叭等关键传感器；下层安装运动控制平台，电池、传感器、转接板和电源板等。

### 2.2.3 电路各模块介绍

为实现高精度自主导航与环境感知，系统构建了集成度高的多源传感系统，涵盖定位感知、环境监测与人体检测。定位模块融合 RPLidar A1 激光雷达、六轴 IMU 与轮式里程计，借助 Cartographer 实现鲁棒的室内定位与建图。其中选用性价比优异的思岚 RPLidar A1 作为主要激光传感器。



图 7 思岚 A1 激光雷达

在环境感知方面，如图 8 所示，机器人搭载 AHT10 温湿度传感器，MQ2 气体，人体红外传感器（PIR）传感器。



图 8 Uboot 用到的传感器

## 2.2.4 转接板与电源板设计

为实现多传感器连接与稳定供电，本系统针对 ELF2 定制扩展板与电源板（图 8），接口板实现 I2C、GPIO 等信号的多通道接入，电源板提供必要的稳压转换与电气保护功能，极大提高了传感系统的可靠性与可维护性。

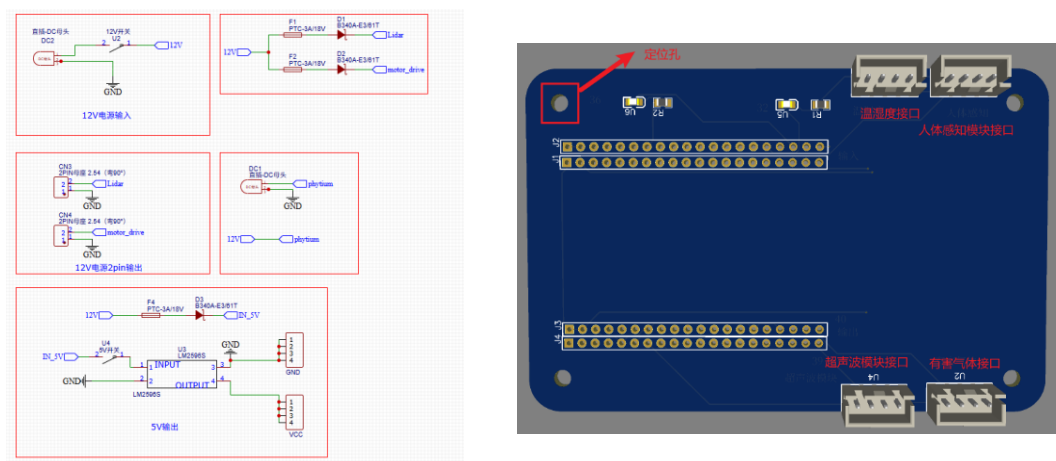


图 9 接口板与电源板

## 2.3 软件系统介绍

### 2.3.1 软件整体介绍

软件系统采用分层架构设计，分为系统层、驱动层与算法层。系统层基于 ELF2 部署裁剪内核的 Ubuntu 20.04，运行 ROS Foxy 系统，负责调度与高层控制；驱动层整合雷达、IMU、里程计、摄像头、音频模块及多类执行器，支持底层硬件控制；算法层基于 ROS 2，集成 Cartographer 建图、Navigation2 导航、语音交互与目标检测等模块，实现复杂任务的智能响应与情感互动。

如图 10 所示，系统分为语音模块，定位模块，规控模块，传感器驱动和感知模块。系统通过 **ROS2 调度机制**合理划分任务，构建了语音交互、导航控制与



环境感知协同运行的软硬件体系。

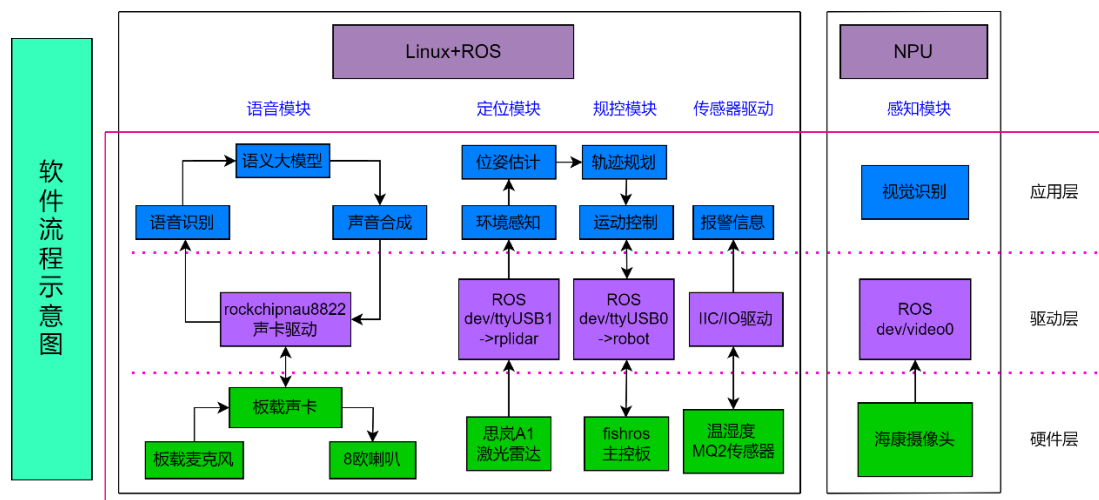


图 10 软件架构

其中。语音模块通过板载声卡接入麦克风，本地 Vosk 模型 实现语音识别，调用讯飞 Spark 4.0 Ultra API 接口实现语义理解，调用百度 REST 实现语音合成，进而实现播报反馈。

导航与定位模块，运行位姿估计、路径规划。基于 Nav2 框架与 Cartographer 进行建图与定位，结合 DWA 实现局部避障，控制指令通过 cmd\_vel 驱动底盘。同时负责系统状态监测，通过 ROS2 与其他模块通信，实现实时信息处理。

NPU 主要负责高性能视觉识别任务，承担图像中目标物体的实时检测工作。系统通过 video0 接口接入摄像头图像流，预处理后由 ROS 节点将图像数据传输至 NPU 推理引擎，加载轻量化 YOLOv5-Lite 模型进行目标识别。能够充分发挥 NPU 算力，实现高帧率、低延迟的视觉感知能力。

### 2.3.2 软件各模块介绍

本项目中的酒店服务机器人主要承担迎宾引导、语音交互、环境监测与智能播报等任务。针对以上功能，我们设计了任务调度器节点。

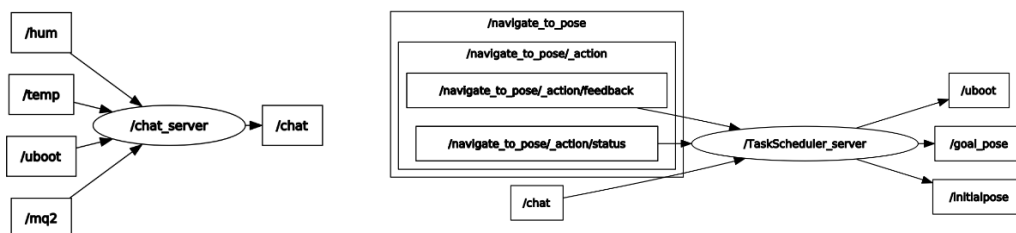


图 11 任务调度器节点图

如图 11 所示，/TaskScheduler\_server 是酒店服务机器人系统的核心调度节点，负责统一协调导航、语音交互、环境监测与状态控制等任务。它通过与 /navigate\_to\_pose、/chat、/goal\_pose、/uboot 等节点交互，接收来自环境传感器（如 /hum、/temp、/mq2）和语音模块的数据，调度机器人执行迎宾引导、语音回复、状态播报等功能，实现多源信息融合与任务流程的有序控制。

### (1) 基于 Vosk 的实时语音识别

为提升酒店服务机器人的本地语音识别能力，项目在 ELF2 平台上部署了轻量级 Vosk 语音识别系统。Vosk 是基于 Kaldi 框架的开源语音识别工具，具备跨平台、离线运行、低延迟等优势，适合在资源受限的国产嵌入式平台上部署。

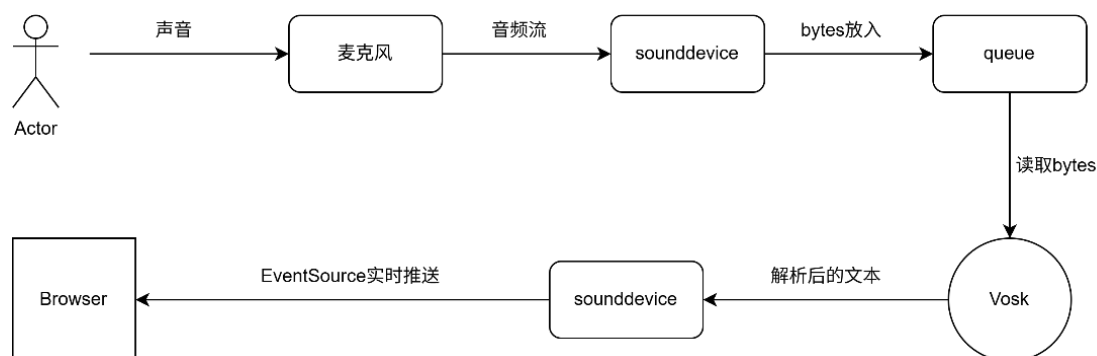


图 12 Vosk 模型结构图

本项目采用 vosk-model-small-cn-0.22 中文轻量模型（如图 12 所示，模型体积仅约 **42MB**，部署于 ROS 2 系统中。系统通过 USB 麦克风采集语音信号，经 Vosk 实时处理后输出识别文本，响应时间控制在 3 秒以内。

### (2) 基于 Vosk 的实时语音识别基于 Spark 4.0 与 REST 的语音交互系统设计

本项目基于 Spark4.0 与 REST 接口的语音对话系统。Vosk 模型负责本地语音转文字，识别文本由设定的 calibration agent（Spark4.0 Ultra 引擎）先进行校正，然后再由 robot agent（Spark4.0 Ultra 引擎）进行分析，命中“导航”“温湿度”等功能词时**触发回调**，非指令文本通过百度 REST API 进行自然语言补全与情感化播报。

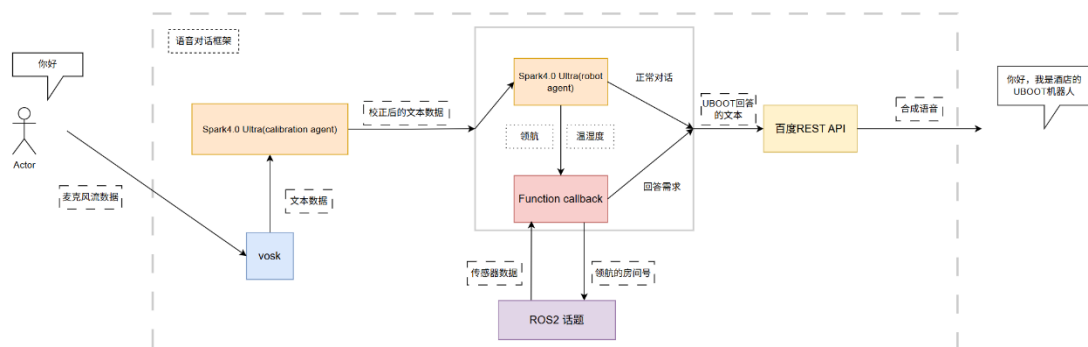


图 13 语音对话系统

在系统融合方面，语音模块通过 ROS 2 的话题机制与导航、环境感知等节点深度协同，总体架构如图 13 所示。

### (3) YOLOv5-Lite 模型部署优化

为赋予智能酒店服务机器人高效的视觉检测能力，本项目部署轻量级目标检测算法 YOLOv5-Lite，实现多类物体实时检测。

该模型结构在 YOLOv5 基础上进行了优化，采用 SFB（Shuffle Feature Block）轻量化模块替代部分 CSP 结构，同时引入通道重排（Channel Shuffle）、深度可分离卷积（DWB）等机制，尤其适配 ELF2 嵌入式平台。

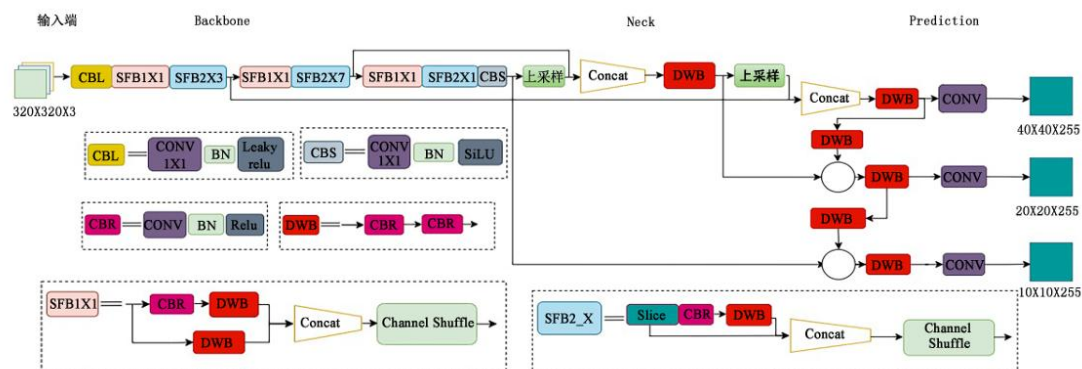


图 14 YOLOv5-Lite 结构图

采用 2564 张训练图像与 640 张测试图像构建定制数据集，涵盖纸巾、瓶子、药品、塑料、口罩、烟雾六类目标。图像分辨率为 320×320，使用 Adam 优化器与 Cosine 学习率调度，训练 50 个 epoch。最终模型在测试集上取得优异表现：**mAP@0.5 为 0.969、Precision 为 0.922、Recall 为 0.952。**

部署阶段，项目先将训练完成的模型转换为 ONNX 格式，并进一步通过

RKNN Toolkit 转换为 RKNN 格式，以适配 ELF2 平台的 RK3588 NPU 加速推理。推理引擎采用 RKNN 推理架构部署轻量化识别系统，结合自定义后处理模块完成置信度筛选与非极大值抑制（NMS），输出目标类别与置信度评分。系统支持六类目标（如瓶子、烟雾等）的实时检测，部署效果稳定，帧率高，满足嵌入式场景下的边缘视觉识别需求。



图 15 数据集示例

#### (4) 基于 Cartographer 的融合定位建图

为实现高精度、高鲁棒性的室内定位系统，本项目采用了 **Cartographer** 作为核心 SLAM 算法，通过融合激光雷达、IMU 和里程计数据，构建稳定的位姿估计框架。

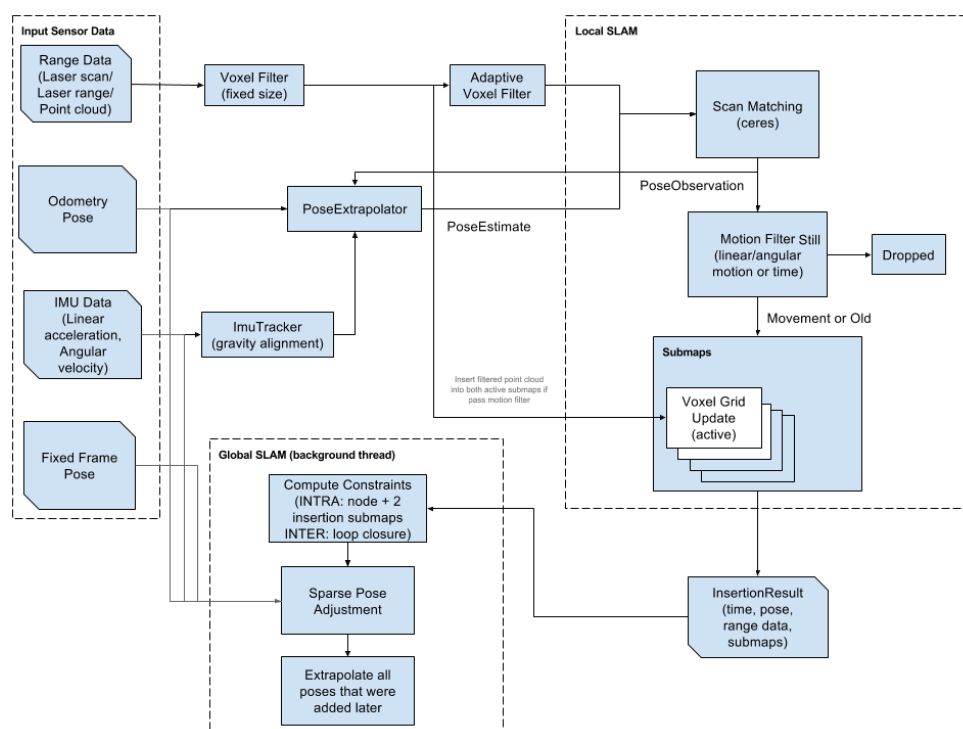


图 16 Cartographer 框架

图 16 中展示的流程包括**本地 SLAM 与全局优化**两个主要阶段：在本地模块中，通过 Scan Matching 与 Motion Filter 实现帧间位姿估计与数据滤波；Submaps 管理模块使用体素栅格地图实时更新。在全局模块中，系统周期性执行回环检测与图优化，降低累计漂移。

## (5) Nav2 自主导航系统设计

本系统基于 ROS 2 Navigation Stack (Nav2) 构建完整的导航流程，如图 17，导航行为由 BT Navigator Server 驱动，通过行为树（Behavior Tree）管理任务逻辑；局部与全局路径分别由 Controller Server 和 Planner Server 生成，底层控制器直接发布/cmd\_vel 速度指令驱动机器人。

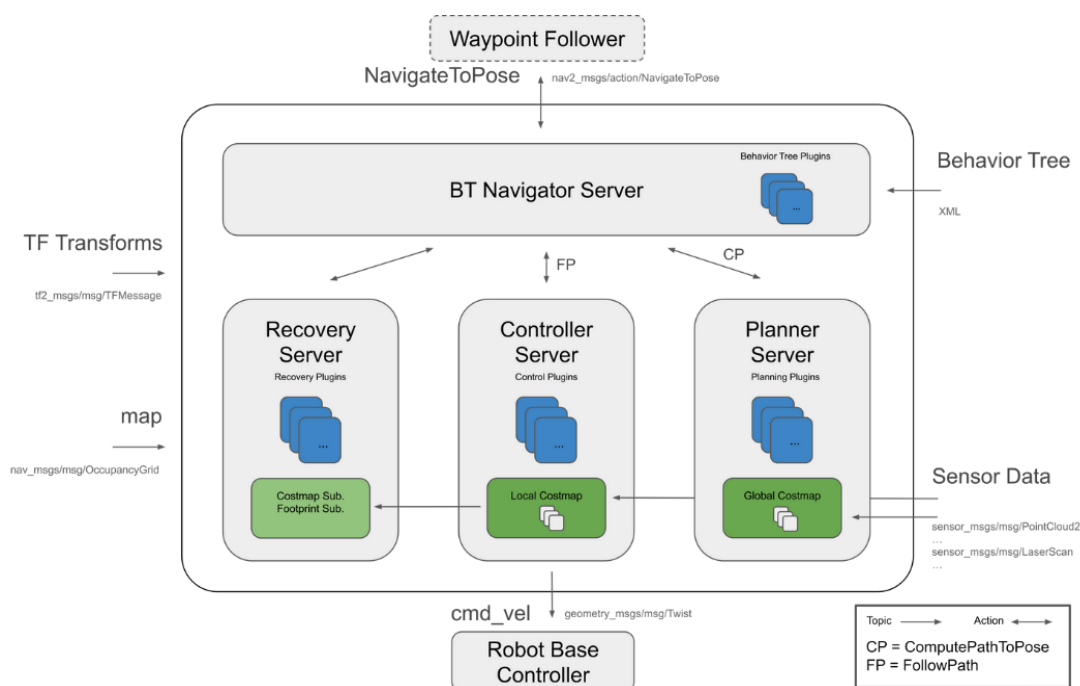


图 17 NAV2 框架图

本项目采用 **DWA 算法**实现局部路径规划，具备良好的动态避障与轨迹优化能力，适应酒店等复杂环境导航。

图 18(见下页) 展示了基于 ROS 2 架构的机器人导航系统节点结构图，包含定位、路径规划、控制与底层通信等模块。

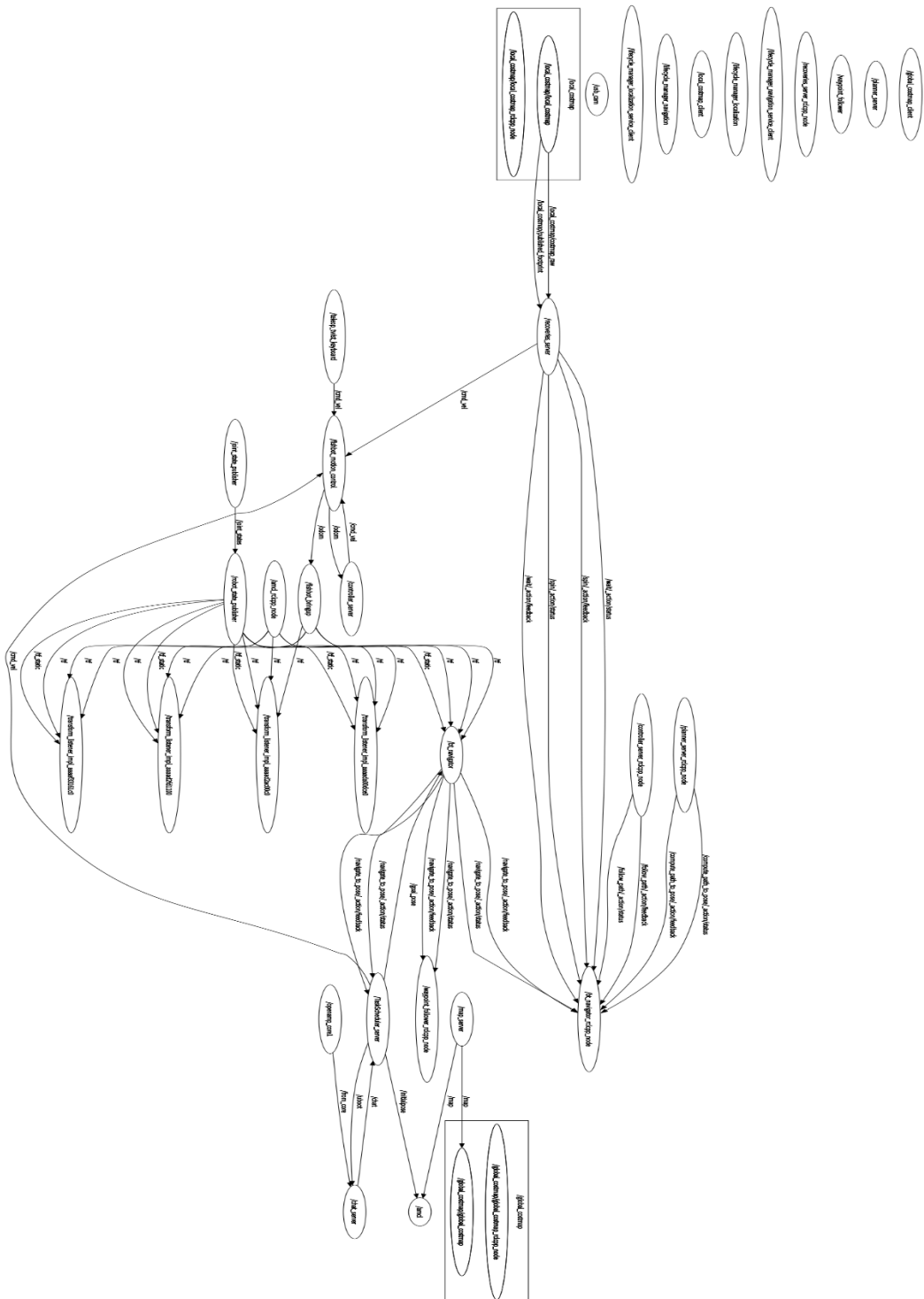


图 18 系统节点图



## 第三部分 完成情况及性能参数

### 3.1 整体介绍

本项目基于 ELF2 平台设计并实现了一套面向酒店场景的智能服务机器人系统，实现了实时定位与建图，路径规划与避障，语音交互与情绪识别，NPU 加速垃圾目标识别，实时环境感知，下图为我们自主搭建的机器人系统。

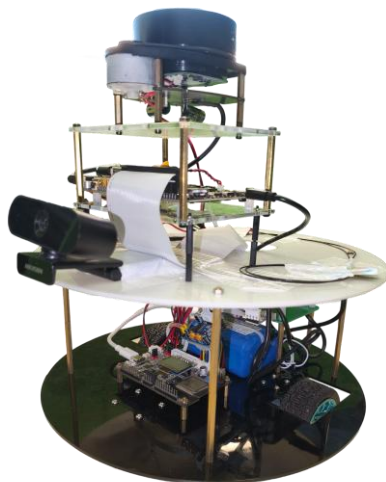


图 19 系统整体图

### 3.2 工程成果

#### 3.2.1 机械成果

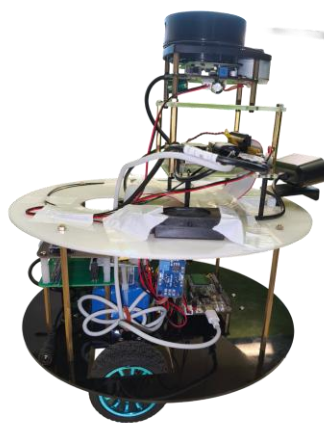


图 20 机械成果展示

#### (1) 分层模块化结构：

机器人采用上下两层平台堆叠设计，上层平台主要安装关键传感器与主控设备，如 ELF2 主控板、激光雷达、摄像头、扬声器等；下层平台用于固定电源模

块、扩展板及轮式底盘，结构清晰，功能隔离，有利于后期维护与调试。

## (2) 高集成度传感装配：

顶部搭载一体式 360°激光雷达模块，用于实现 SLAM 定位与建图；前方装有广角摄像头，支持目标识别与交互检测。中部装有麦克风与语音播报装置，支撑语音交互功能。

## (3) 差速轮式底盘：

底部采用两主动轮+一组万向轮的经典差速驱动结构，结构紧凑，转弯灵活，适应室内复杂空间。主动轮由底层电机驱动，结合编码器实现精确里程计反馈；万向轮提升平衡性。

### 3.2.2 电路成果

如图 6，图 9（第 8 页）所示，我们自主设计了电源板、转接板等核心电路，满足系统集成与稳定运行需求，体现出良好的工程化与集成能力。

### 3.2.3 软件成果

本项目集成了，SLAM，NAV，QT 上位机，语音模块，目标检测模块，传感器驱动模块。

图 21 为 SLAM 模块的建图以及定位效果。

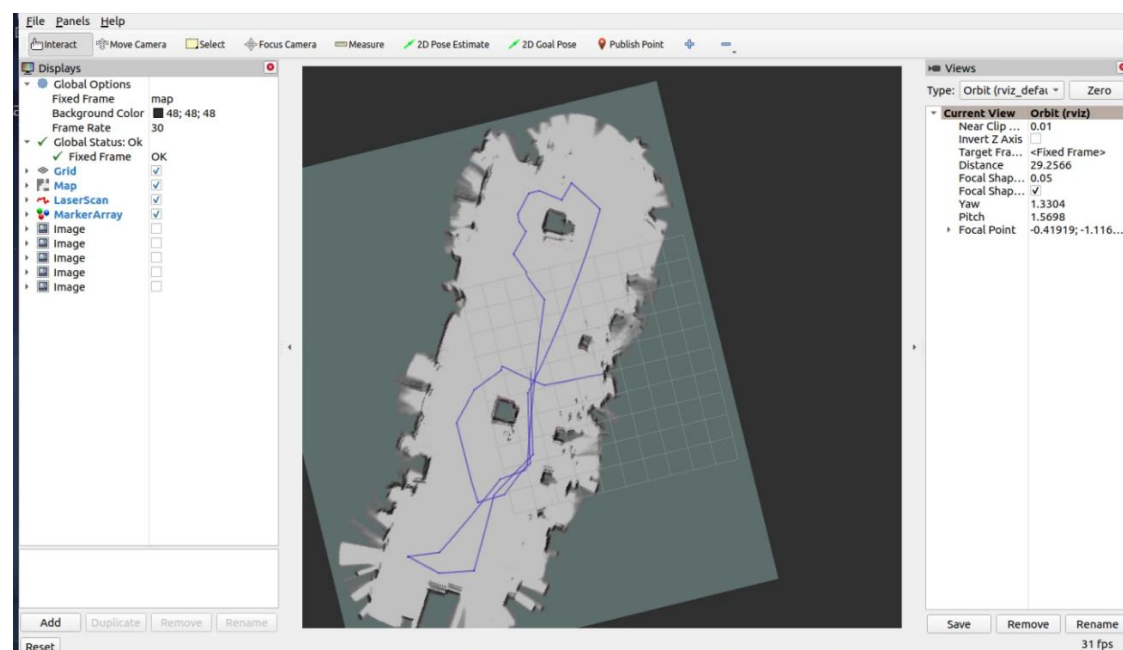


图 21 定位与建图

图 22 展示了路径规划与控制的动态功能：

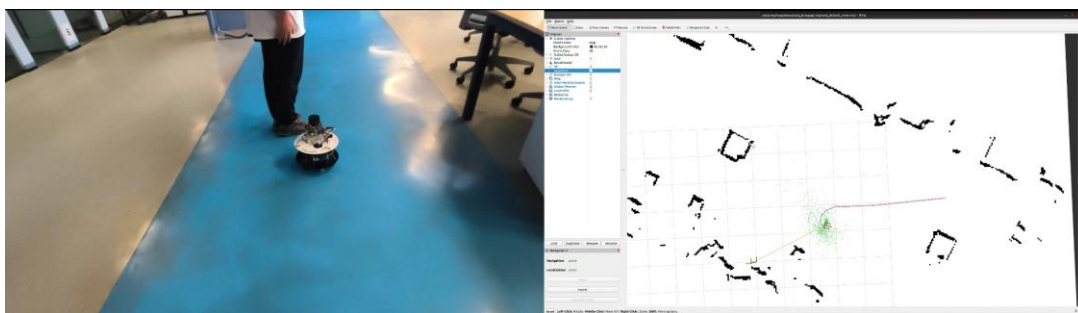


图 22 避障功能

图 23 展示了上位机的 UI 接入与环境信息动态显示。

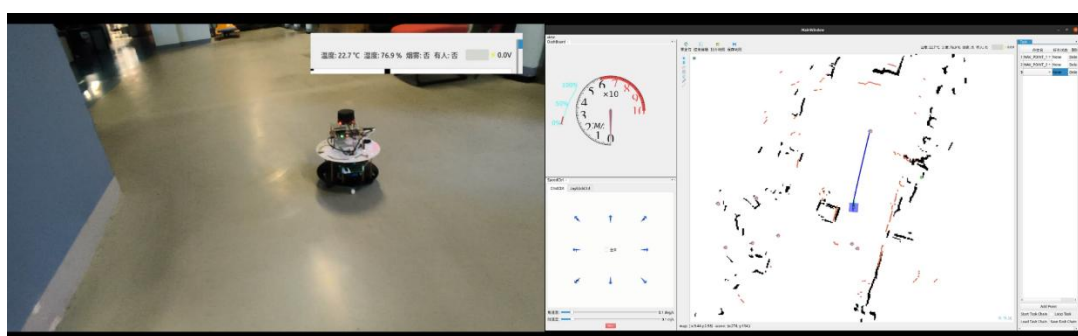


图 23 上位机 UI

图 24 展示了目标检测的实时识别。

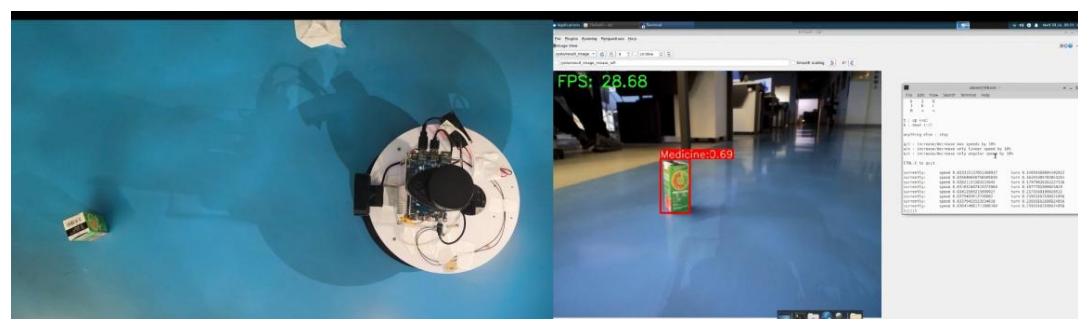


图 24 视觉检测模块

图 25 展示了使用语音大模型进行对话与环境信息获取。

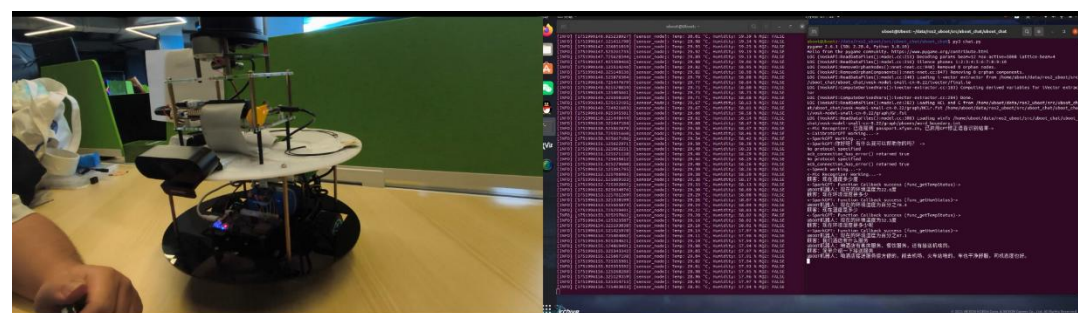


图 25 大模型对话与传感器驱动

图 26 是在酒店场景进行导航的场景实录。

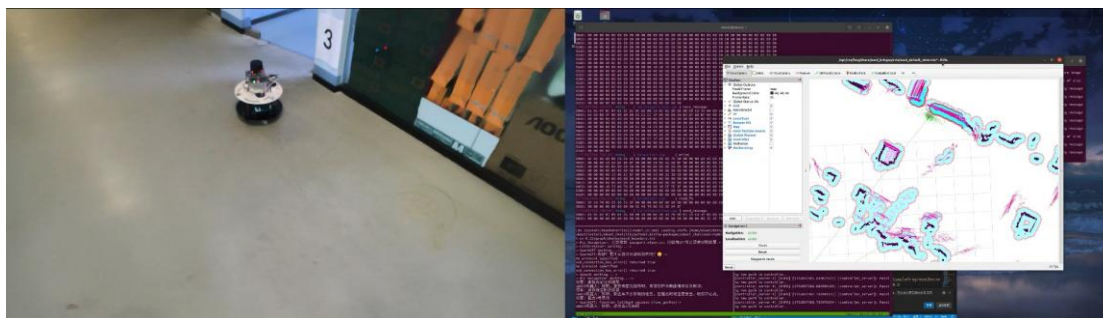


图 26 酒店领航功能展示

### 3.3 特性成果

#### (1) 建图功能

在较长走廊模拟酒店内部环境。操作人员控制机器人行进，实时生成地图与定位。测试中进行转弯、回环与遮挡等场景，评估系统的建图连续性与鲁棒性。

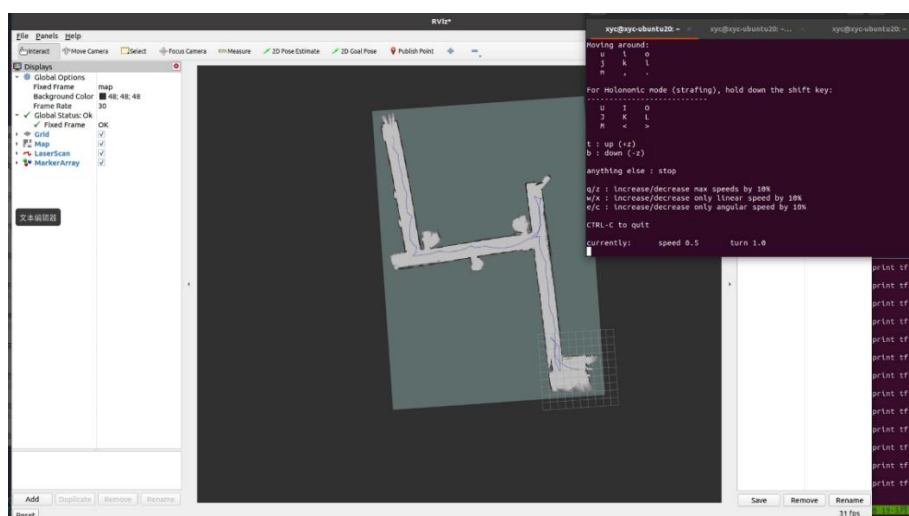


图 27 建图测试

对应的系统指标如下：

表 2 建图指标

参数	指标
地图几何精度误差	$\pm 50\text{ cm}$
建图效果	无明显跳帧、漂移、崩图
测试场景	大于 $50\text{ m}^2$ 的复杂空间建图良好

将建图系统在 100 平方米左右的长廊退化场景里进行建图测试，下图为在 100 平方米模拟场景建立的 2D 栅格地图：

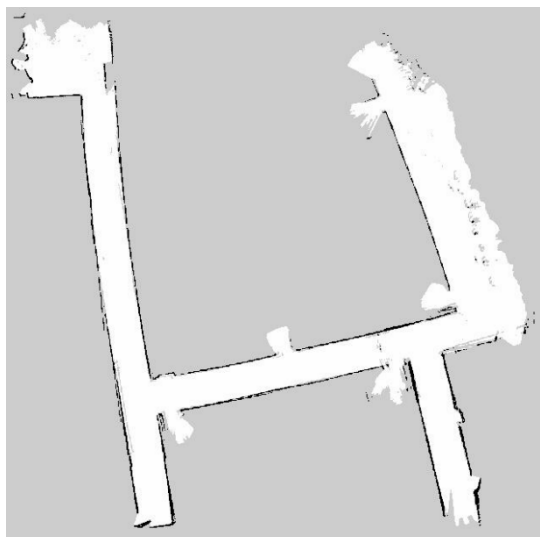


图 28 模拟酒店地图

由上图可见系统成功完成了连续建图，地图边缘闭合良好，结构清晰，表明算法的实用性与稳定性。

## (2) 导航功能

从起点发起多段导航指令（前台→房间→前台），系统自动规划避障前进。包括人员干扰、障碍物动态出现等情况。

系统的性能指标如下：

表 3 导航指标

参数	指标
到点成功率	$\geq 95\%$
局部避障成功率	$\geq 90\%$
平均路径规划时间	$< 1000\text{ ms}$
轨迹偏移误差	$\leq 15\text{ cm}$ （室内）

如图 30 所示，系统在 ELF2 平台上成功运行 Nav2 导航模块，地图结构完整、边界清晰，激光点云与静态轮廓高度一致，反映出良好的定位效果；规划轨迹平滑，表明导航系统运行稳定。

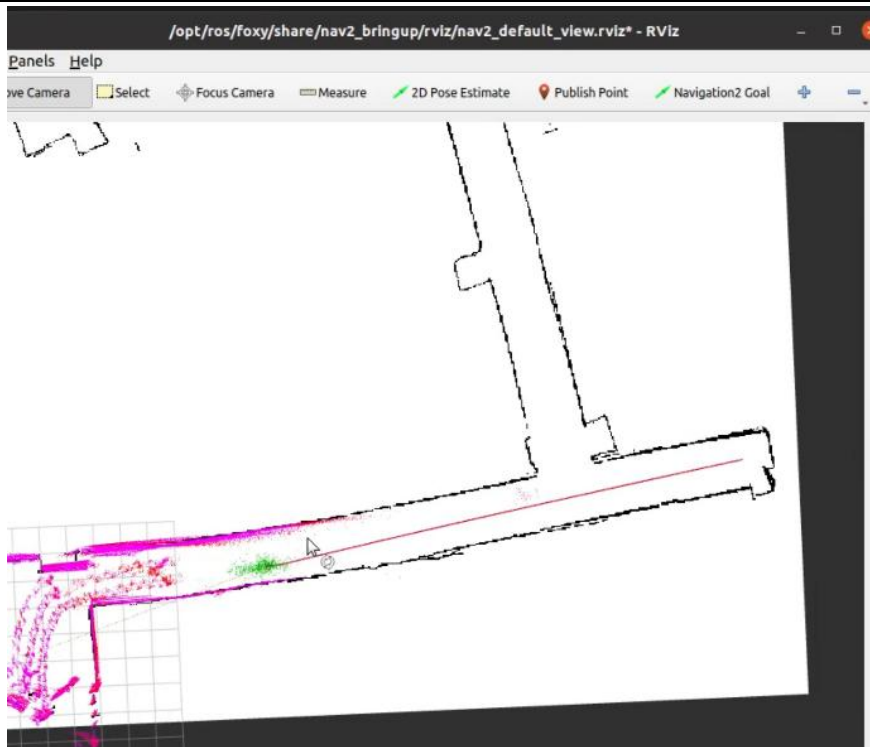


图 29 大场景导航测试

图 30 为机器人的多点导航功能展示。

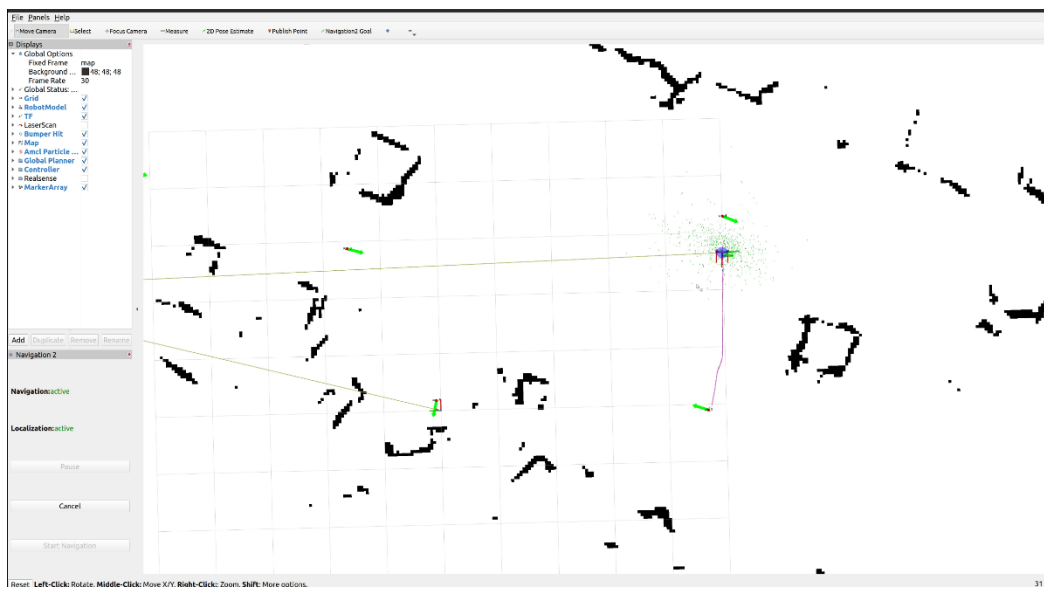


图 30 多点导航

### (3) UI 测试

系统移植并修改了开源的远程 UI 项目。具有选择目标点、遥控、查看任务状态、环境信息等功能。

系统指标如下：



表 4 UI 指标

参数	指标
按钮响应时间	小于 300ms
功能入口跳转成功率	100%
切换刷新	无明显卡顿或闪屏现象
刷新周期	$\leq 200\text{ms}$

图 31 为测试过程中的 UI 界面截图：

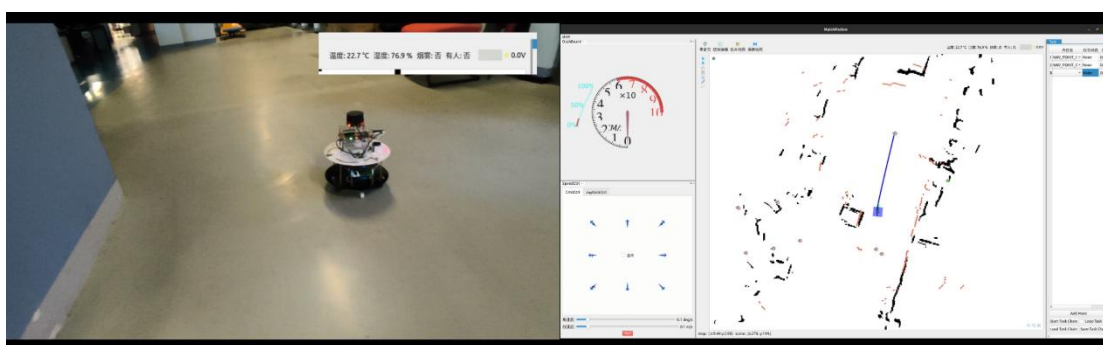


图 31 UI 界面

显示了酒店机器人 UI 在实际部署过程中的操作界面。通过图片效果可见 UI 效果正常。整体界面简洁、功能齐全，适合酒店等复杂场景中的机器人导航与调度任务。

#### (4) 目标检测

测试场景模拟真实酒店环境，考察算法的识别效果。

在 ELF2 测试效果如图 32 所示，能够准确识别出多类常见物品：纸巾、塑料、口罩、药品。并附带置信度得分与类别标签。

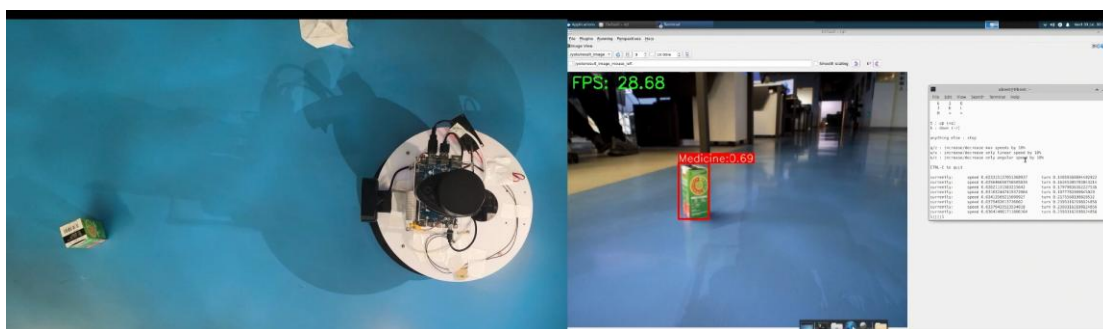


图 32 目标检测效果

模型速度达到平均 25~30 FPS 的实时检测能力，能满足大部分服务动作前

的前置感知需求。

在实际场景测试中，机器人能够稳定检测出地面散落的纸巾、塑料片、口罩等物体，表 5 展示各类物体在真实环境中的平均识别效果：

表 5 目标检测效果

类别	Precision	Recall	稳定性评价（备注）
纸巾（Tissue）	90.3%	92.5%	表面反光略影响边界，识别稳定
瓶子（Bottle）	95.6%	96.4%	表面清晰，对背景不敏感
药品（Medicine）	94.2%	95.1%	包装印刷清晰，鲁棒性较好
塑料（Plastic）	91.8%	93%	易与纸类混淆，需增强细粒度识别
口罩（Facemask）	81.5%	90.2%	受折叠影响，易发生遮挡
烟雾（Smoke）	85.4%	91%	可检测轻度烟雾，环境光影响
平均值（ALL）	89.8%	93%	具备实用级性能，满足部署需求

从表格可以看出，系统在真实环境中的 Precision 与 Recall 略低于标准测试集，但整体仍保持在较高水平，尤其在瓶子与药品等具备明确边缘与纹理的目标上识别效果稳定。

相比之下，软质目标（如口罩、纸巾）在部分姿态或光照条件下存在一定误差，仍可接受识别率；烟检测则对光线较敏感。视觉检测系统整体在真实场景下具备良好的实时性、准确性与稳定性。

(5) 语音识别

用户依次输入常见指令（如“你好”、“带我去 3 号房间”、“今天温度是多少？”），测试在普通环境和噪声环境下的识别效果。

实际测试后，本项目的指标达到如下效果：

表 6 语音测试

项目	指标要求
识别准确率	≥ 85%
响应延迟	≤ 3 秒
抗噪稳定性	识别率下降 ≤30%
连续对话成功率	≥ 90%

测试效果如下图所示，由图可见语音识别准确，语义对话顺畅：

```
shock/vosk-model-shock-ckpt/gpt/phones/normal-phones/phones
<-Mic Recognizer: 已连接到 passport.xfyun.cn, 已启用GPT修正语音识别结果->
<-CalibratorGPT Working...->
<-SparkGPT Working...->
<-SparkGPT:你好呀!有什么我可以帮你的吗? ->
No protocol specified
xcb_connection_has_error() returned true
No protocol specified
xcb_connection_has_error() returned true
<-Speech working...->
<-Mic Recognizer Working...->
顾客: 现在温度多少度
<-SparkGPT: Function Callback success (func_getTempStatus)->
UBOOT机器人: 现在的环境温度为22.6度
顾客: 现在环境湿度是多少
<-SparkGPT: Function Callback success (func_getHumStatus)->
UBOOT机器人: 现在的环境湿度为百分之78.4
顾客: 现在温度是多少
<-SparkGPT: Function Callback success (func_getTempStatus)->
UBOOT机器人: 现在的环境温度为32.3度
顾客: 现在环境湿度是多少呢
<-SparkGPT: Function Callback success (func_getHumStatus)->
UBOOT机器人: 现在的环境湿度为百分之87.1
顾客: 我们酒店有什么服务
UBOOT机器人: 咱酒店有客房服务、餐饮服务, 还有接送机啥的。
顾客: 简单介绍一下接送服务
UBOOT机器人: 咱酒店接送服务挺方便的, 能去机场、火车站啥的, 车也干净舒服, 司机态度也好。
顾客: 我能带你去你的房间
```

图 33 语音识别

## (6) 系统测试

用户通过语音指令要求机器人引导前往酒店的 3 号房间，测试系统对自然语音的识别、指令解析、路径规划与导航执行能力。

下表为我们达到的系统指标展示

表 7 系统指标

测试项目	指标要求
语音识别准确率	$\geq 85\%$
指令响应延迟	$\leq 1.5\text{ s}$
导航路径规划时间	$\leq 1000\text{ ms}$
任务执行成功率	$\geq 90\%$
导航轨迹偏移误差	$\leq 15\text{ cm}$ （室内环境）

图 37 展示了机器人在用户发出，“请带我去 9 号房间”语音指令后的完整响应流程。系统先通过 Vosk 模型进行本地语音转写，随后调用 Spark 4.0 完成语义解析，成功触发导航任务（func\_goto）。整个过程连续流畅，语音识别、指令解析、导航控制环节协同顺畅。

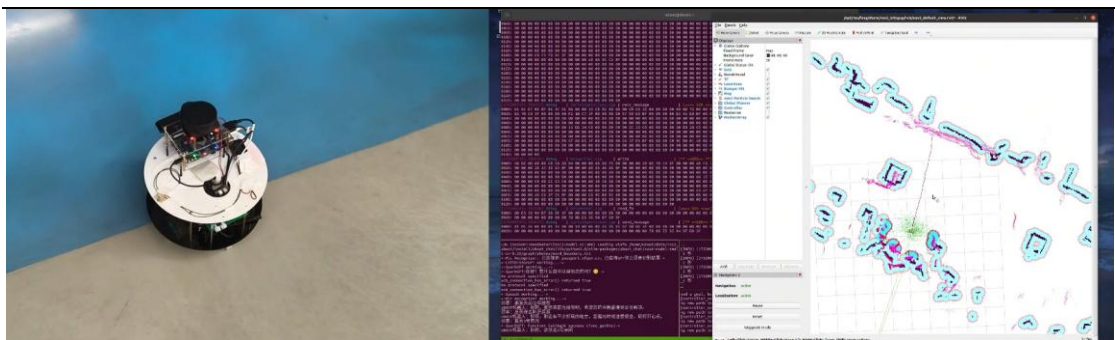


图 34 系统测试效果

该场景体现了语音识别、语义理解、导航规划与运动控制等模块的高效集成，充分展示出系统在复杂任务执行中的模块融合能力与运行稳定性。

## 第四部分 总结

本项目基于 ELF2 平台，成功实现了一套面向酒店场景的智能服务机器人系统。完成了以下工作：

实现融合激光雷达、IMU 与里程计的 SLAM 定位建图，部署 ROS 2 Nav2 完成路径规划与避障；集成讯飞与百度语音 API 实现语音交互及情绪识别；基于 NPU 使用 YOLOv5-Lite 进行垃圾目标识别；通过 FishROS 控制板完成底盘控制，并自主设计电源与转接板；定制化机械结构适配酒店场景，优化传感器安装与维护；完成高集成度系统 UI 与整体部署，并通过多场景测试验证系统稳定性。

### 4.1 可扩展之处

尽管系统整体运行稳定，但仍存在需优化的问题：语音识别在嘈杂环境中存在误差，情绪识别能力有限；目标检测在低光或强反光条件下精度下降；建图与定位对传感器精度依赖较高，复杂动态环境中仍有轻微位姿漂移；UI 界面受限于网络与设备性能，存在访问延迟与刷新不及时现象。

为提升系统整体性能与应用价值，未来将从以下方向持续优化：加强语音交互智能化，在本地 NPU 部署 LLM，提升语音对话的响应速度和稳定性；升级视觉感知，融合 RGB-D 等多模态数据提升检测鲁棒性；优化 Cartographer 与 Nav 配置，增强动态环境下定位与导航稳定性；完善与酒店平台的数据接口，实现任务调度与状态反馈的自动化闭环管理；开展多机器人协同研究，基于 ROS 2 DDS 构建分布式服务体系；搭建云端学习平台，支持模型迭代与行为优化，增强系统自适应与规模部署能力。

### 4.2 心得体会

在本次“UBoot”智能酒店服务机器人项目的研发过程中，我们深切体会到了“理论与实践融合”的重要性。本项目以 RK3588 国产芯片为核心平台，融合了多模态感知、异构算力调度与智能交互技术，从立项到测试部署，经历了完整的系统工程流程。

首先在**架构设计**阶段，我们确立了“CPU 调度控制 + NPU 视觉推理”的异构协同机制。这不仅大幅提升了计算资源的利用效率，还实现了语音交互与目标检测等任务的解耦并行。在硬件方面，团队自主设计了电源板与转接板，使得多



种传感器（激光雷达、IMU、气体传感器等）能够稳定集成。尤其是在转接板布线、稳压保护和接口规范等细节中，我们逐步摸索出一套适配国产平台的稳定解决方案。

在**软件系统**开发中，语音模块采用本地 Vosk 与 Spark 4.0 云服务协同机制，实现了离线转写、语义识别与情感反馈的闭环交互。通过对接 ROS 2 框架与 REST 接口，系统能在 3 秒内完成语音识别与响应，达到了良好的实时交互体验。同时，我们针对语音指令中“房间导航”类的指令词进行了定制解析，大大提升了识别精度与用户体验。

**视觉模块**方面，我们训练并优化了轻量级 YOLOv5-Lite 模型，结合 RKNN Toolkit 部署于 NPU。在模型压缩过程中，我们引入通道重排（Channel Shuffle）与深度可分离卷积（DWB）等策略，使其在不牺牲精度的前提下，将帧率提升至 28–30 FPS，mAP@0.5 高达 96.9%。在实际场景中，系统能够稳定识别纸巾、口罩、烟雾等六类物体，为后续的垃圾检测与语音反馈提供了有力支持。

**导航部分**是系统稳定性的关键。通过融合 RPLidar A1 激光雷达、IMU 与轮式里程计，系统基于 Cartographer 实现高鲁棒 SLAM；路径规划采用 Nav2 与 DWA 算法，测试中轨迹偏移控制在 15 cm 以内。在酒店等动态环境中仍能完成多点导航、路径重规划与障碍绕避，展现出良好的实用性。

整个项目中，我们深刻认识到团队协作和工程调优的重要性。从底盘结构设计到 SLAM 参数调试、再到语音模型响应延迟的优化，每一个细节都决定了系统的整体体验。尤其是在多模块集成过程中，调试不同节点通信频率、内存占用和线程阻塞成为关键挑战，最终我们通过合理调度 ROS 2 节点、细化任务优先级与频率配比，成功实现了高集成度与系统稳定性。

此次研发不仅加深了我们对机器人系统工程的理解，也验证了国产硬件平台在智慧服务领域的强大潜力。我们相信，随着后续在语义理解、多轮对话、云端训练与多机协同等方面的拓展，UBoot 有望在实际酒店场景中实现更广泛的应用。



## 第五部分 参考文献

- [1] Yuan B, Su J J, Dai C L. Research on Visual Recognition of Service Robots Based on YOLO Algorithm[J]. *Journal of Computing and Electronic Information Management*, 2024, 9(28).
- [2] 宋海荣, 张媛媛, 杨少魁. 智能机器人语音交互专利技术分析[J]. *河南科技*, 2020(01).
- [3] 陈燕, 王佳盛, 曾泽钦. 大视场下荔枝采摘机器人的视觉预定位方法[J]. *农业工程学报*, 2019(01).
- [4] 李志翔, 黄剑华, 甘仿. 基于 RK3588 的老人智能监控系统设计[J]. *软件工程与应用*, 2024(01).
- [5] Liang H X, Li Y, Guo Q Q. ROS2-based locator optimized autonomous navigation robot[C]//Proceedings of the 2023 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI). 2023.
- [6] Zhang Z Y, Li R W, Wang Y. Research on Voice Interaction Robot System Based on the Hybrid Chat Model of Retrieval and Generation[J]. *International Journal of Advanced Computer Science and Applications*, 2023, 8(4).
- [7] 李红光, 马永桃, 蒋晨曦. 基于 ROS 的驱鸟机器人控制系统研究[J]. *科技与创新*, 2024(01).
- [8] 丁鑫, 谢婷婷, 巴梓轩. 基于瑞芯微平台的轻量化人脸属性估计视频流设计方法[J]. *电光系统*, 2024(01).
- [9] Gao Y C, Gao Z Q, Chen X H. Robot Vision Recognition System Based on Improved YOLOv3 Algorithm[J]. *International Journal of Robotics and Automation*, 2023, 3(4).
- [10] Sun Y, Wang X, Chen Z, et al. An Efficient and Robust Visual SLAM System for Indoor Environments[J]. *Robotics and Autonomous Systems*, 2023, 164: 104238.
- [11] Liu H, Wang X, Zhang Y, et al. Real-time and Accurate Speech Recognition for Service Robots in Noisy Environments[J]. *Applied Acoustics*, 2024, 213: 108836.
- [12] Wang Y, Li Z, Liu Q, et al. A Lightweight Object Detection Algorithm for Robot

---

Vision with High-speed and High-precision[J]. *Journal of Visual Communication and Image Representation*, 2024, 93: 104598.

[13] Zhao X, Qian X, Zhou Y, et al. Multimodal Sensor Fusion for Autonomous Navigation of Service Robots in Complex Indoor Environments[J]. *Sensors*, 2024, 24(10): 4678.

[14] Chen S, Liu F, Wu X, et al. An Adaptive Path Planning Algorithm for Service Robots in Dynamic Environments[J]. *Control Engineering Practice*, 2024, 145: 105773.

[15] Li Q, Zhang X, Wang Y, et al. A Novel Semantic Parsing Method for Voice Interaction of Service Robots[J]. *Information Sciences*, 2024, 680: 1-15.

[16] Zhang H, Sun D, Li C, et al. Improving the Robustness of Visual Object Detection for Service Robots under Different Lighting Conditions[J]. *Pattern Recognition Letters*, 2024, 181: 1-8.

[17] Wang L, Liu X, Zhang J, et al. An Integrated System for Service Robots Combining Navigation, Voice Interaction and Visual Detection[J]. *Journal of Intelligent & Robotic Systems*, 2024, 107(3–4): 1–18.

[18] Radxa Docs Team. RKNN Quick Example on RK3588[EB/OL]. (2024-09-20)[2025-07-09].

[19] Feng D. object\_detection\_tracking\_rk3588: 基于视频输入的二维目标检测与跟踪示例[EB/OL]. (2024-12-02)[2025-07-09].