

叶政韩,吕文涛,郭庆,等. 基于树莓派的 AI 识别智能车控制系统[J]. 智能计算机与应用,2025,15(2):47-51. DOI:10.20169/j. issn. 2095-2163. 250207

# 基于树莓派的 AI 识别智能车控制系统

叶政韩<sup>1</sup>, 吕文涛<sup>1</sup>, 郭庆<sup>2</sup>, 徐羽贞<sup>2</sup>, 余润泽<sup>3</sup>

(1 浙江理工大学 信息科学与工程学院, 杭州 310018; 2 浙江省技术创新服务中心, 杭州 310007;

3 中国移动通信集团设计院有限公司浙江分公司, 杭州 310012)

**摘要:**智能车在世界范围的应用已经从航空航天领域转向日常生活中,针对目前智能车控制方式过于单一的问题,本文提出了一种基于树莓派的智能车控制系统,通过集成人脸、语音和手势,实现多种智能交互方式,并结合百度智能云平台 and Tornado 框架实现高效的数据处理与远程控制。系统采用百度智能云平台的识别算法处理人脸、语音和手势数据,确保高精度识别;远程控制模块基于 Tornado 框架,通过 Web 编辑器实现树莓派本地端对远程 Web 端的实时监听;识别模块和控制模块协同工作,保证实时数据反馈与智能车的驱动控制。系统成功实现了基于树莓派的 AI 智能车控制,能够实时准确地识别输入的数据,并通过 Tornado 框架实现远程端与本地端的实时通信与控制。数据反馈被转换为智能车的驱动控制信号,完成了从数据采集到智能控制的全流程。

**关键词:** 百度智能云; Tornado 框架; 树莓派; 智能车

中图分类号: TP181

文献标志码: A

文章编号: 2095-2163(2025)02-0047-05

## AI recognition intelligent vehicle control system based on Raspberry Pi

YE Zhenghan<sup>1</sup>, LV Wentao<sup>1</sup>, GUO Qing<sup>2</sup>, XU Yuzhen<sup>2</sup>, YU Runze<sup>3</sup>

(1 School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China;

2 Zhejiang Province Technology Innovation Service Center, Hangzhou 310007, China;

3 China Mobile Communications Group Design Institute Co., Ltd. Zhejiang Branch, Hangzhou 310012, China)

**Abstract:** The application of intelligent vehicles in the world has slowly shifted from the aerospace field to daily life. In this thesis, an intelligent vehicle control method based on Raspberry Pi is studied, and three control methods are realized: face control, voice control, gesture control and network remote control. Three different kinds of recognition modules invoke the recognition algorithm of Baidu Intelligent Cloud Platform to achieve high-precision recognition of face, gesture and voice data. The remote-control module establishes a small server through the web editor based on the Tornado framework, which realizes real-time monitoring from the local side of the Raspberry Pi to the remote web terminal. This system finally realizes the research and design of AI intelligent vehicle control system based on Raspberry Pi, completes the real-time accurate identification of the collected data by the intelligent cloud platform and the real-time monitoring of the local end of the Raspberry Pi based on the Tornado framework, and outputs the data feedback in the mode of intelligent vehicle drive control.

**Key words:** Baidu Intelligent Cloud; Tornado frame; Raspberry Pi; intelligent vehicle

## 0 引言

近期有关智能小车的议题,如坡道行驶电动小车、智能送药小车以及跟随小车等,智能小车各种功能的开发设计也越发丰富<sup>[1-3]</sup>。同时智能云的发展也迎来了高潮,通过使用智能云平台,可以大大降低

硬件负担以及硬件成本。百度智能云平台的智能识别功能强大,识别正确率高,广泛运用于各种场景,为了应对越来越高的硬件成本问题,其平台的技术水平也在不断发展<sup>[4]</sup>。树莓派作为一种低成本、高性能的单板计算机,已成为智能硬件开发中的重要平台,为智能小车的研发提供了强有力的支持。

**基金项目:** 国家自然科学基金(U1709219,61601410); 浙江省科技厅重点研发计划(2021C01047,2022C0079); 2021年产业技术基础公共服务平台项目(2021-0174-1-1)。

**作者简介:** 叶政韩(2001—),男,硕士研究生,主要研究方向:工业互联网和数字图像处理。

**通信作者:** 吕文涛(1982—),男,博士,副教授,主要研究方向:工业互联网和数字图像。Email:alvinlwt@zstu.edu.cn。

收稿日期: 2023-07-26

本文设计并实现一个基于树莓派的智能小车控制系统,融合了人脸识别、语音识别和手势识别等多种智能交互方式,以提高智能车的操控多样性和智能化水平。通过集成百度智能云平台提供的识别算法和 Tornado 框架,实现了高效数据处理和远程控制。人脸、语音、手势 3 种输入方式与树莓派控制系统的结合,不仅提升了系统的识别精度,也使得用户可以通过直观的方式与智能小车进行交互。本系统不仅可以扩展智能车控制的方法和途径,还为未来智能硬件的发展和应用提供了有益的探索和参考。

## 1 硬件选型与设计

本文所用到的开发板为树莓派 4B,其性能比其它型号的树莓派低,使得该平台特别适合作为比较和评估智能云识别与本地算法识别在实际应用中的性能差异,特别是其对树莓派资源占用的影响。

作为智能车的驱动控制模块,本设计采用了 L298N 电机驱动模块。该电机驱动控制模块的输出电压与输入电压有关,电机驱动模块硬件资源配置如图 1 所示,其中电机 A 和 B 为智能车左侧电机,C 和 D 为右侧电机。

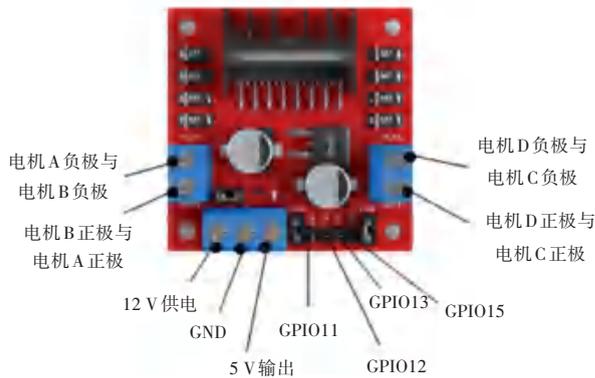


图 1 L298N 资源配置图

Fig. 1 Resource configuration diagram of L298N

树莓派摄像头模组采用 800 万像素,支持 1 080p30,720p60 和 VGA90 录像模式的模块,图像回传效果较好,延迟较低,且与其他摄像头相比,延迟表现较为出色,非常适合作为图像识别的数据采集模块<sup>[5]</sup>。蓝牙音箱搭载的是蓝牙 5.3 版本,具有高于 85 dB 的信噪比和小于等于 0.5% 的失真率,整体功能和参数均超出最低要求。

在电机驱动模块的连接与配置中,将树莓派与电机驱动控制模块 L298N 连接,通过其输出端与小车的驱动电机相连。PWM (Pulse-Width Modulation)

应用脉宽调制进行模拟电路控制,其信号输入端被连接至树莓派的使能端,负责输出控制脉冲信号,以便调节电机的转速<sup>[6]</sup>。L298N 的控制逻辑关系见表 1,进一步明确了各控制信号与电机运行状态之间的对应关系。

表 1 L298N 控制逻辑关系表

Table 1 Logical control relationship of L298N

直流电机	旋转方式	IN1	IN2	IN3	IN4	ENA	ENB
A1、A2	正转	高	低			高	
	反转	低	高			高	
	停止	低	低			高	
A3、A4	正转			高	低		高
	反转			低	高		高
	停止			低	低		高

硬件设备整合后的整体设计框图如图 2 所示。

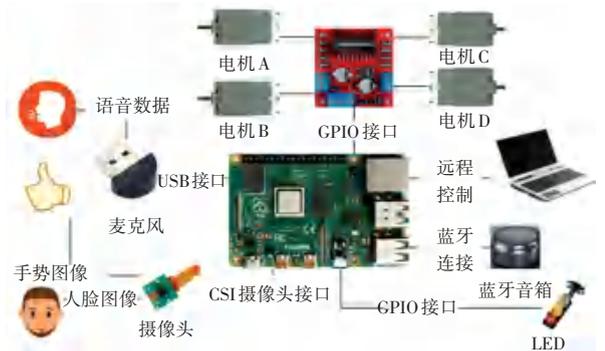


图 2 硬件设计整体框图

Fig. 2 Design drawing for hardware

## 2 识别控制系统的设计与实现

识别控制系统建立在百度智能云平台上,系统主要包括 3 个控制途径:人脸识别、手势识别和语音识别。

### 2.1 人脸识别

#### 2.1.1 图像数据采集与预处理

图像预处理方面,由于在网络上进行数据传输时,服务器对数据的处理方式不尽相同,容易造成数据在传输过程中出现偏差,造成数据的丢失所以根据智能平台的需求,需要将人脸图像数据进行 Base64 编码<sup>[7]</sup>。Base64 编码是指将一副图片的数据编码成一套字符串,使用字符串来代替图像地址。Base64 编码将不规范的字符转换为符合 URL 规范的字符,经过处理后数据接收端即可正常接收数据<sup>[8]</sup>。图像编码前和编码后示例如图 3 和图 4 所示。



图 3 编码前图像示例

Fig. 3 Example of pre-coding image

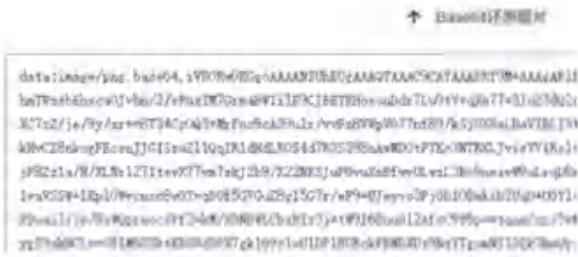


图 4 编码后图像

Fig. 4 Example of encoded image

### 2.1.2 人脸识别功能的实现

人脸识别功能的实现首先需要完成鉴权认证机制在百度 AI 智能云平台注册账号,首先在智能云平台创建人脸识别、语音识别和肢体识别功能,获取 API Key 和 Secret Key,在官方平台领取免费的各个功能识别的额度,该额度足够日常的使用。用户可以通过调用百度 AI 的 SDK (Software Development Kit) 库中的 AipFace 类,根据提供的 API 参数发送请求,从而获取认证令牌 (Access Token)。一旦获取到该令牌,用户即可获得平台上各种智能识别服务的完整使用权限,每个功能模块,如人脸识别、语音识别等都对应一个特定的端口,在实际应用中,用户需要为每个功能模块单独申请额度,并通过获取对应端口的权限来实现具体的操作。

完成鉴权认证后,查找智能云平台的人脸识别系统功能。首先,在智能云平台上创建相关的人脸识别库,百度智能云会生成专属 ID 以及密钥;在库中录入需要进行认证的人脸,多个人脸可以分为不同或者相同的用户组如图 5 所示;在用户组中添加用户并命名如图 6 所示,在对应用户下导入对应人脸 (生活照) 即可完成一个人脸信息的录入,同时对于已经录入的人脸可以进行导入和导出操作来更换用户组,更好的分类管理。

为了实现与云平台的有效对接,需要在树莓派上安装相应的环境。利用 HTTP SDK 为树莓派与云端平台之间建立通信桥梁,从而实现数据的交互和

人脸识别功能的远程调用。通过该 SDK 的支持,树莓派能够通过网络与百度智能云平台实时连接,确保人脸识别服务的顺利运行。本系统通过树莓派摄像头实时获取图像,并将图像数据编码为 Base64 格式,以便上传至百度智能云进行人脸识别;智能云平台对上传的图像进行处理后,返回相关识别结果,包括是否检测到人脸、识别到的身份信息 (如人名) 以及人脸相似度等数据。树莓派按照设定的时间间隔周期性拍摄图像,并将每次拍摄的图像上传至智能云平台。系统会在每次识别后根据返回结果判断是否继续拍摄,直到成功识别出正确的人脸为止。

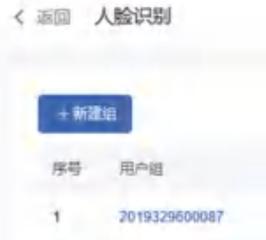


图 5 人脸用户组

Fig. 5 Face user group

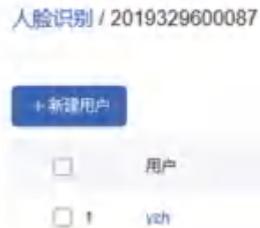


图 6 人脸用户

Fig. 6 Face user

为了实现系统识别过程的可视化,本文通过本地显示实时拍摄的图像并以幻灯片形式播放,同时在界面上输出关键数据,如图像上传时间、识别结果及相似度等信息,用户能够直观地观察系统的识别过程和效果,为后续的优化与调试提供数据支持。实时图像数据如图 7 所示。



图 7 实时图像数据

Fig. 7 Real-time image data

人脸识别完成后,百度智能云会返回各种图像详细数据,其中最重要的数据分别为人脸的唯一标

识和人脸置信度。人脸标识为对人脸进行身份确认;人脸置信度为测试人脸识别准确率,人脸信息主要的返回参数见表2。

表2 人脸信息主要的返回参数

Table 2 Main return parameters of face information

字段	类型	说明
face_num	int	检测到图片中人脸数量
face_list	array	人脸信息列表
+face_token	string	人脸图片的唯一标识(有效期60 min)
+location	array	人脸在图片中的位置
++left	double	人脸区域离左边界的距离
++top	double	人脸区域离上边界的距离
++width	double	人脸区域的宽度
++height	double	人脸区域的高度
++rotation	int64	人脸框相对于竖直方向的顺时针旋转角
+face_probability	double	人脸置信度(0~1)
...	...	...

## 2.2 手势识别

图像采集与预处理具体步骤与人脸识别所使用的图像处理方法基本一致。

识别图片中的手势类型,返回手势名称、手势矩形框、概率分数,可识别24种常见手势,适用于手势特效、智能家居手势交互等场景。

图片中有多个手势时也能正确识别,但该情况下单个手势的目标可能较小,且存在角度因素影响,如存在倾斜、遮挡等,识别效果可能有偏差。针对单个手势识别时效果最佳。手势识别主要返回参数见表3。

表3 手势识别主要的返回参数

Table 3 Main return parameters of gesture recognition

字段	类型	说明
result_num	int	数量结果
result	object[ ]	检测到的目标
+classname	string	目标所属类别
+top	int	目标框上坐标
+width	int	目标框的宽
+left	int	目标框最左坐标
+height	int	目标框的高
+probability	float	目标属于该类别的概率
log_id	int64	唯一的 log_id,用于问题定位

## 2.3 语音识别

### 2.3.1 音频数据采集与预处理

音频文件处理方面,调用 pyaudio 库进行音频录制与处理,利用 pygame 库进行音频播放。利用导入的 pyaudio 库接收音频数据,并将数据保存为 wav 格式

(不是唯一合适的格式)。保存录音文件需要配置多种合适的参数,根据智能云平台需要的格式进行转换。

### 2.3.2 语音识别功能的实现

语音识别以及语音播报使用的智能识别功能为同一个 ID 以及密钥。通过调用百度 AI 的 SDK 库中的 AipSpeech 类,并根据提供的 API 参数发送请求,从而获取 Access Token。

首先,需对语音数据进行采样与预处理,以确保数据质量适合后续的识别任务;其次,将处理后的音频文件(通常为 WAV 格式)上传至智能云平台进行语音识别。为提高识别的准确性与针对性,系统要求设置相应的语言识别代码,以标识待识别的语种。不同的语言和方言对应着不同的语言代码,如普通话(1536)、纯中文普通话(1537)、英语(1737)、粤语(1637)以及四川话(1837)等。

语音识别完成,系统会返回多种形式的输出结果,通常包括语音内容的文本转录以及可能涉及的语音特征,如语音语调等。这些输出数据可为进一步的语音分析提供基础,通过对特定语音特征,如语调、语气等的深入分析,可实现对特定语言风格或情感的识别与分类。

在数据传输和交互方面,智能云平台通常采用标准的 JSON 格式进行数据的串行化,以便进行高效的传输与解析。识别任务的相关数据,包括语音文件及其对应的参数配置,通常以 JSON 格式通过 POST 请求进行上传,此种数据格式具有良好的扩展性和通用性,可以适应不同识别任务和参数配置的需求。JSON 数据结构中包含了必要的参数字段,具体内容可依据具体应用需求进行定制,并通过平台的接口进行数据传递与处理。JSON 包含的参数见表4。

表4 JSON 所包含的参数

Table 4 Parameters of JSON

字段	类型	说明
format	string	语音文件格式
rate	int	采样率
channel	int	声道数
cuid	string	用户唯一标识
token	string	开放平台获取到的开发者
dev_pid	int	不填写 lan 参数生效,默认 1537
lm_id	int	训练平台模型 id
lan	string	历史兼容参数
speech	string	本地语音文件的二进制语音数据
len	int	本地语音文件的字节数

### 3 远程控制系统的设计与实现

智能车的远程控制系统基于 Web 界面的设计,其中前端界面布局和按钮响应依赖于 Tornado 框架。Tornado 是一种高效的 Web 服务器框架,其非阻塞式设计使其能够在高并发的情况下保持较高的响应速度,相比于当前主流的 Web 服务器框架,Tornado 框架在处理大量并发请求时表现出明显的优势,这对本系统的实时性要求至关重要。

本远程控制系统的核心功能是本地端与远端之间的数据传输与交互。在本地端构建一个小型 Web 服务器,远端通过网络连接该服务器,并将控制指令以数据包的形式传输给本地端;接收到数据后,本地端对数据进行解析,并通过控制树莓派的 GPIO 口输出相应的控制信号,从而驱动小车的硬件,实现了远程命令的接收与反馈过程,通过网络实时的远程控制智能车。

#### 3.1 本地端设计

##### 3.1.1 电机驱动模式设计

本文设计的智能车的方向控制相对简单,无法通过姿态传感器等设备进行精细的姿态调整,因此,姿态控制的精度主要依赖于修改代码中的延时参数,并结合对智能车运动状态的观察来进行调整。尽管这种方法可以在一定程度上改善智能车的运动效果,但由于缺乏精确的姿态反馈机制,本系统无法实现如 X/Y 轴定位等精确的导航控制功能。

小车运动模式的设计依照表 1 的 L298N 电机驱动控制模块逻辑功能表进行编码。

##### 3.1.2 对远端的监听设计

树莓派的编译程序通过 Tornado 框架实现对远端命令行指令的监听与处理。使用 Tornado 的 options 函数库解析命令行参数,定义、转换并存储全局参数。通过这种方式可以将端口号作为参数传入,避免端口号被硬编码在程序中,使程序在运行时能够灵活地根据库函数来解析传入的参数。

建立 Tornado 实例后,所有导入的变量会传递到 Tornado 应用类,与 Tornado 库进行数据交换,确保 Tornado 库能够正确调用相应的处理程序(handlers)。handlers 通常是一个元组列表,用于定义不同 URL 路径对应的处理逻辑<sup>[9]</sup>。进一步地,通过 Tornado 的 HTTPServer 对象,可以指定端口并实时接收命令行数据;指定端口可以实时获取来自远端的数据;通过 Tornado 的 IOloop 实例,便开始接收并处理 HTTP 请求,确保远程控制指令能够及时

响应<sup>[10]</sup>。

#### 3.2 远端设计

远端控制的编码主要通过 HTML 文件实现。在按钮响应设计中,采用键盘事件(keydown、keyup等)来捕获用户的输入行为,并在事件触发时记录对应的命令行响应。

为了确保操作的即时性,只需要按下按钮而无需松开即可触发事件,在用户按下按钮时立即执行相应的动作,而无需等待释放操作,前一个事件执行完毕后,才会触发下一个事件。

为了实现清晰且易于操作的用户交互,界面被划分为多个区块,每个区块都赋予了明确的名称和功能。

### 4 结束语

本文设计并实现了基于树莓派的智能车控制系统,人脸、语音和手势识别与智能车的控制功能相结合,显著提升了智能车的交互性和智能化水平。通过集成百度智能云平台的识别算法和 Tornado 框架,实现了高效的数据处理、精确的识别效果以及远程实时控制能力。系统不仅能够精确识别用户输入的信息,还能够通过实时数据反馈与控制信号的转换,确保智能车驱动的稳定性与响应性。该基于树莓派的智能控制系统在多模态交互和远程控制领域具有巨大潜力,为未来智能车的应用发展提供了新的思路与技术支持。

#### 参考文献

- [1] 孙弋婷. 基于 MSP430 单片机的坡道行驶电动小车设计[J]. 电子制作, 2021, 423(13): 10-11.
- [2] 吴必端, 魏榛旭, 姜良麒. OpenMV 智能送药机器人小车的设计与实现[J]. 宁德师范学院学报(自然科学版), 2023, 35(1): 52-56.
- [3] 张春梅, 温新权, 胡高天, 等. 智能跟随小车的设计与制作[J]. 机电产品开发与创新, 2022, 35(5): 75-78.
- [4] 李志勇. 百度智能云多层面自主创新推进实体经济智能升级[N]. 经济参考报, 2022-10-19(008).
- [5] 刘泽伟. 基于树莓派的自动避障小车的设计与实现[J]. 电子制作, 2022, 30(17): 78-80.
- [6] 张秉森, 马吉忠, 杨一飞, 等. 基于树莓派的自动避障小车的设计与实现[J]. 南方农机, 2021, 52(3): 23-24.
- [7] 周岚. 一种基于 Base64 编码的多图片上传技术研究[J]. 电脑知识与技术, 2018, 14(31): 106-108.
- [8] 陈禹航. 私有 base64 编码转化分析及其 Python 实现[J]. 电脑与电信, 2022, 311(12): 87-92.
- [9] 张也. 基于声明式交互的数据可视化系统的设计与实现[D]. 杭州: 浙江大学, 2018.
- [10] 刘涛, 孙茜, 易善凯. 基于 Tornado 的课程管理系统的设计与实现[J]. 软件工程, 2018, 21(12): 22-24.