

---

## 基于 RK3588 的大语言模型辅助驾驶系统

### 摘要

本作品基于 RK3588 开发板构建大语言模型辅助驾驶系统,通过融合 ResNet 与 LSTM 构建轻量化道路预测模型,实现道路状态实时检测(畅通/缓行/拥堵/封闭),结合多序列投票策略提升稳定性,准确率达 90%,推理速度 0.2 秒。检测结果同步触发 RGB 灯光提示(如红/黄/绿/蓝对应路况)与语音安全播报,降低无反馈辅助驾驶风险。

创新集成 Deepseek-R1 大模型与 RAG 框架,本地化部署地图数据库。用户可通过语音模块实时交互导航请求,模型响应时间约 40 秒,支持全离线运行。

系统具备低功耗与高实时性,突破传统方案在复杂路况适应性和交互灵活性上的局限,并可通过更新 RAG 数据库扩展至特种车辆场景(如冷链运输),为车载智能系统提供安全预警与自然语言导航的创新解决方案。

---

## 第一部分 作品概述

### 1.1 功能与特性

本作品聚焦于交通状况识别和车内语音提示两个重要的车辆辅助驾驶功能，通过计算机视觉技术完成对道路状况的检测，以板卡 LED 模仿车内灯光提示，引入 Deepseek-R1 大模型，将道路检测状态实时通过语音进行反馈。同时构建 RAG 模型框架，构建道路导航数据库及知识数据库，通过语音模块完成用户与 Deepseek 的实时沟通，给出更为精准且互动性更强的信息提示。

与传统车辆辅助驾驶系统相比，该作品放弃使用雷达等传统方式，改为计算机视觉技术对道路状态的检测更为快速，且具有一定的预测性，可以应对市内环境等更加复杂的路况减少急加速/急刹车，降低能耗、驾驶安全增强，且可作为车机系统内的一部分与灯光进行联动，更好提示司机，避免长期依赖无反馈辅助驾驶带来的安全隐患。

传统导航功能无法实现用户与导航系统的实时沟通，本作品通过引入带 RAG 框架的 Deepseek 大语言模型来实现用户与导航系统间的实时沟通，可以给予更加便捷的提示，同时可以实现离线部署，完成特殊情况下的导航功能。

### 1.2 应用领域

本作品的应用领域聚焦于车内辅助驾驶系统，完成道路状态检测和语音提示，可广泛应用于家用车车机系统，同时可以修改 RAG 模型框架下的数据库，应用于工业等领域增加专业的用车提示，如冷冻运输行业冷藏箱温度设置等提示。

本作品设计离线和在线两种模式，在线模式下语音输出更加流畅，而在离线模式下作品仍保留核心功能，既可面向正常情况下的车机系统，更加适用于极端情况下如无人区断网等的车辆辅助驾驶。

### 1.3 主要技术特点

**可扩展性：**在 RAG 模型框架下，数据库内容可以根据具体应用场景实时更新，既可用于导航系统也可以增加工业用车方面的专业知识给予用户精准的提示反馈。

易用性：模型预测与 Deepseek 资源消耗总体较小，能够满足大多数的算力限制要求，同时全系统可应用于离线状态下。

高效性：与传统方法相比，通过计算机视觉对道路进行预测可以达到帧的级别，在我们的算法下，对道路状态的预测准确度可达 90%，更新时间可以控制在 0.8 秒左右。

稳定性：通过摄像头实时采集图像数据，可以避免信号干扰等情况对传统方法的影响。

## 1.4 主要性能指标

作品主要性能指标涉及预测模型速度、精度和 Deepseek 响应速度，旨在为用户提供快速、准确且便捷的服务。

表 1 性能指标图

评判对象	指标	分析
语音识别响应速度	约 0.5 秒	可以满足实时性要求
Deepseek 响应速度	约 40 秒	可以满足实时性要求
道路检测模型推理速度	约 0.2 秒	可以满足实时性要求
道路检测模型识别准确性	准确率约为 90%	集成策略有效降低误判率
道路检测模型资源占用	模型大小约为 32.7MB 显存占用小	适合嵌入式部署
Deepseek 资源占用	内存占用约 2.5GB	适合嵌入式部署

## 1.5 主要创新点

### (1) 语音交互与 Deepseek-R1 实时沟通

作品部署 Deepseek-R1，通过语音及音频模组让用户可直接与 deepsee 对话沟通，搭建 RAG 模型框架，为用户提供更加准确的信息。

### (2) 计算机视觉技术应用于实时道路检测，更加精准稳定

该作品在软件方面，设计独特的道路预测模型，打破传统方式稳定性不强的限制，且模型资源消耗少适用于嵌入式部署。

### (3) 可全离线部署，应用前景广泛

Deepseek-R1 与预测模型皆部署于开发板内部，且可以根据具体使用场景要

求，更改数据库内容，实现功能精准定位。

## 1.6 设计流程

### 1.6.1 需求分析

目前市面上辅助驾驶由于缺乏对快速处理道路信息的能力，在用户过于依赖其时极容易发生安全事故，并且目前的导航系统不能够实现司机的实时询问，容易造成误解等问题，因此我们准备设计一款大语言模型辅助驾驶系统。

### 1.6.2 架构流程

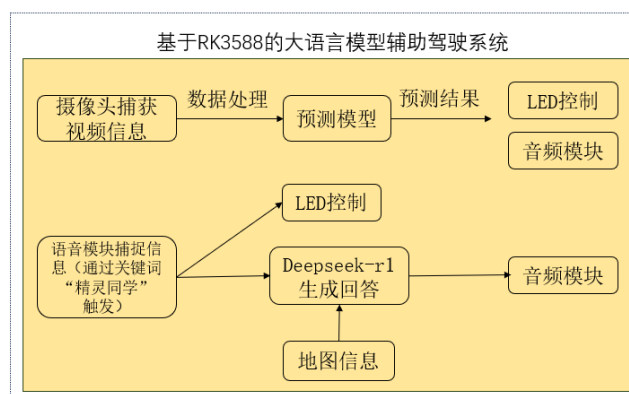


图 1 系统设计架构图

### 1.6.3 问题修复

在离线状态下，语音合成输出效果较差，考虑到正常使用情况下联网，我们更新了联网情况下的语音模块，让语音合成更加流畅。

## 第二部分 系统组成及功能说明

### 2.1 整体介绍

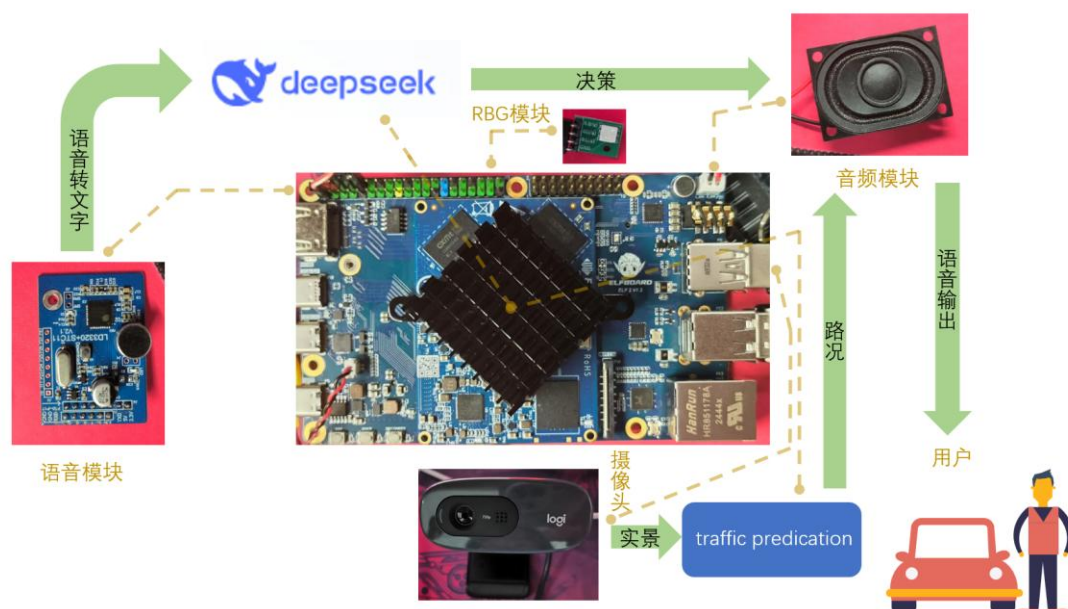


图 2：系统整体框架图

为保证系统可靠性，本作品分为两个流程部分：

**道路预测模组：**将通过训练得到的 traffic predication 模型转为 rknn 格式部署在开发板上，将摄像头信息传入预测模型中进行预测，当预测结果发生变化时会通过音频模块进行语音输出提示，同时 LED 灯会根据预测结果变为不同的颜色，来模拟车内的氛围灯提示。

**Deepseek 对话模组：**在开发板上部署 Deepseek-R1 大语言模型，准备好地图信息库作为 RAG 框架中的数据部分；通过语音模块将用户语音输入的所在地与目的地转化为文字，输入到 RAG 模型中由 Deepseek 根据地图信息生成导航文本再由音频模块输出给用户；

## 2.2 硬件系统介绍

### 2.2.1 摄像头模块

将实时行车画面输入到已经训练好的 traffic predication 中进行判断当前路况



图 3：摄像头模块

### 2.2.2 音频输出模块

输出道路检测安全提示词；将大语言模型所生成的决策，实时播报给用户



图 4：音频模块

### 2.2.3 LD3320 语音识别模块

将用户语音输入转化为数据输入给大语言模型作为提示词。



图 5：语音模块

### 2.2.4 RGB 模块



图 6：灯光模块

提示用户当前运行什么模块以及道路情况，具体如下表：



表 2 灯光控制表

颜色	对应状态
绿色	道路畅通
红色	道路拥堵
黄色	缓行
蓝色	道路封闭
紫色	Deepseek 正在运行

2.2.5 电路各模块介绍

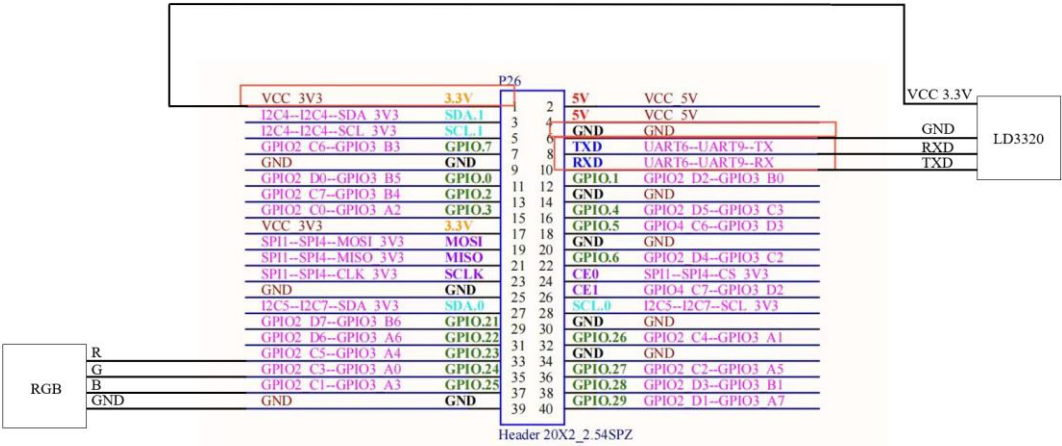


图 7 灯光及语音模块连接图

表 3 LD3320 语音识别模块连线

LD3320	——>	开发板 40pin
VCC 3.3V	——>	VCC 3.3V
TXD	——>	RXD
RXD	——>	TXD
GND	——>	GND



表 4 RGB 模块与开发板连线

RGB	——>	开发板 40pin
R	——>	GPIO.23
G	——>	GPIO.24
B	——>	GPIO.25
GND	——>	GND

## 2.3 软件系统介绍

### 2.3.1 软件整体介绍

软件部分以主函数为驱动，控制预测模块与语音对话模块两部分，通过临时文件的方式控制任务程序交互运行。

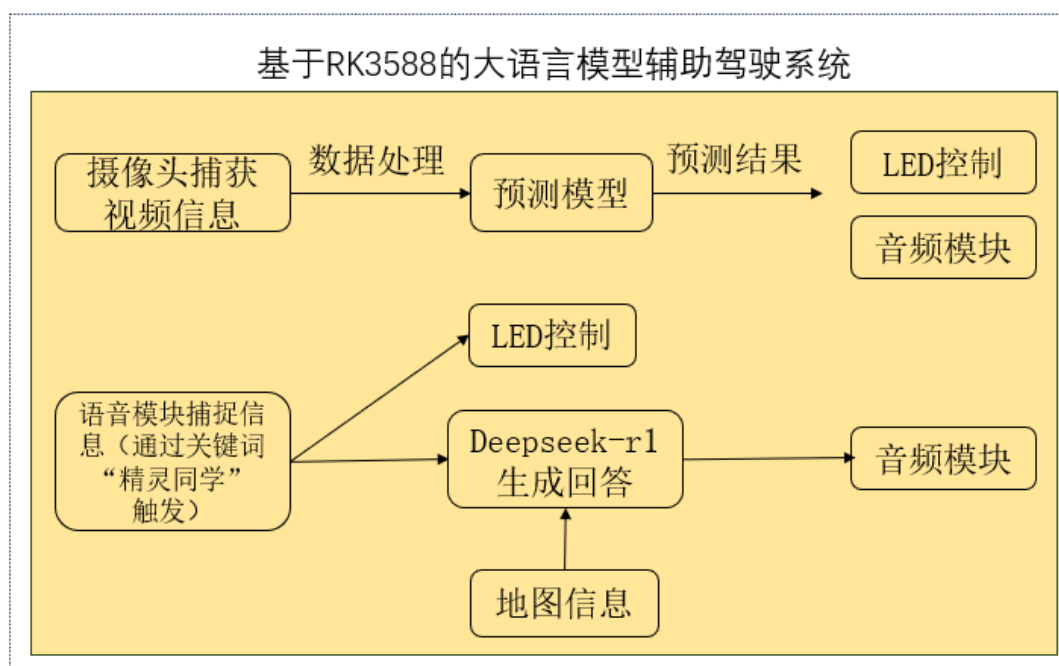


图 8 系统设计架构图

### 2.3.2 软件各模块介绍

模型设计：

在模型设计部分，为达到理想预测效果且满足嵌入式部署对模型资源消耗的限制，本作品以 4 类（畅通、缓行、拥堵、封闭）共计 4288 项、21440 张图片为训练数据，并设计基于 resnet 的预测模型，采取以下策略进行训练：

- 时空融合: CNN 空间特征 + LSTM 时序建模
- 注意力机制: 空间注意力 + 时序注意力双重机制
- 迁移学习: ResNet50 预训练 + 渐进式微调
- 正则化策略: 多层次 Dropout + BatchNorm + LayerNorm
- 温度缩放: 推理时 logits 平滑化，提高预测稳定性

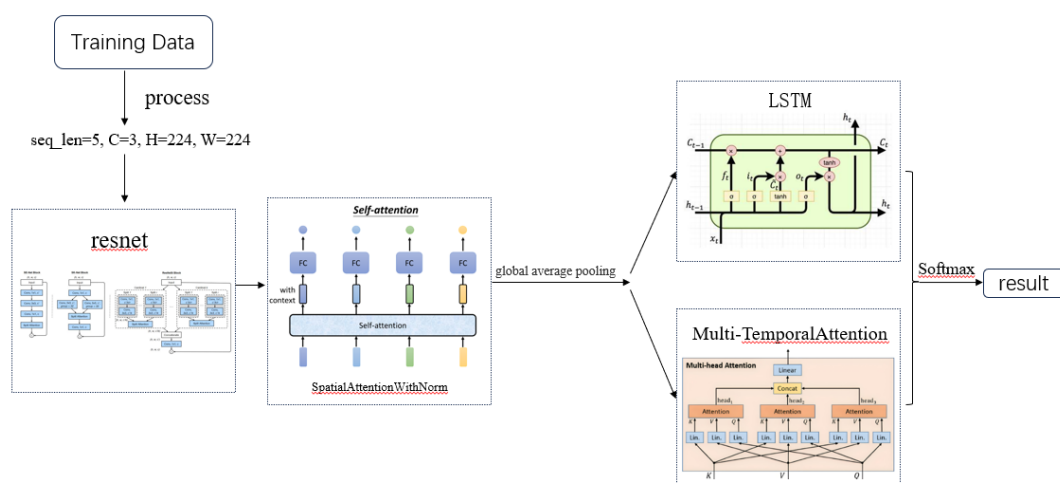


图 9 模型框架图

预测策略：

为提高模型预测精度，作品对预测策略的设计采取了多序列投票、置信度加权的方式进行，同时为避免复杂路况的波动性过大，结果输出部分以 3 次相同为依准确定道路状态变化给出提示。

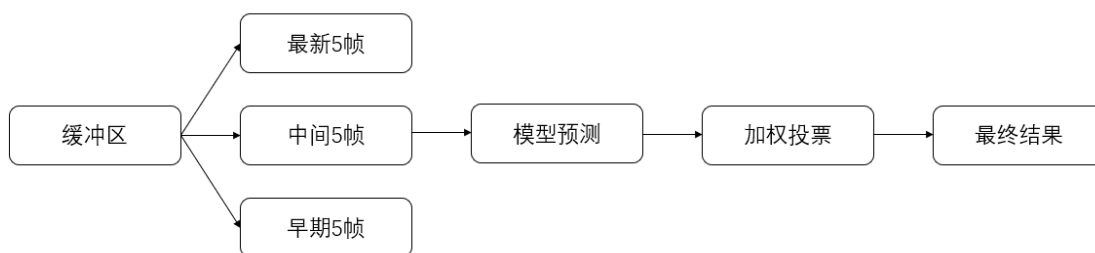


图 10 预测策略图

LED 控制：

表 5 LED 控制图

颜色	对应状态
绿色	道路畅通
红色	道路拥堵
黄色	缓行
蓝色	道路封闭
紫色	Deepseek 正在运行

音频模块：

根据不同使用情况，我们设计了两种音频模块的输出分别为离线状态与在线状态，离线状态通过 `espeak` 命令进行语音合成输出，为提升语音合成质量，特将文本信息转化为拼音形式；而在在线状态下通过微软提供的语音合成系统，语音质量能够有很好的保证。

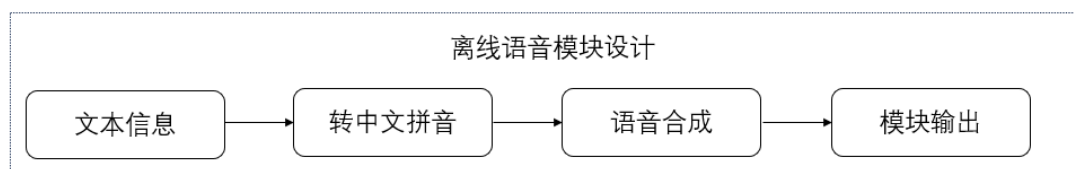


图 11 离线语音模块设计图

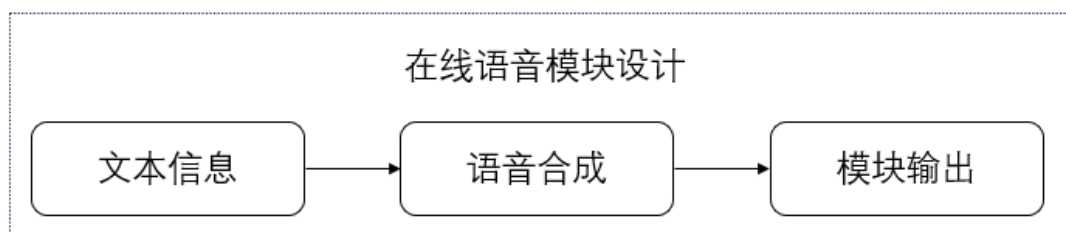


图 12 在线语音模块设计图

以 Deepseek-R1 为生成器构架的 RAG 模型：

RAG 模型即检索增强生成模型，以 Deepseek-R1 为生成器，搭配订制的地图信息库，可以为用户提供较好的信息提示；地图信息提前被转成向量库的形式解读储存，Deepseek 在得到语音模块的信息后，将自身的知识库与数据库信息结合输出有用信息。

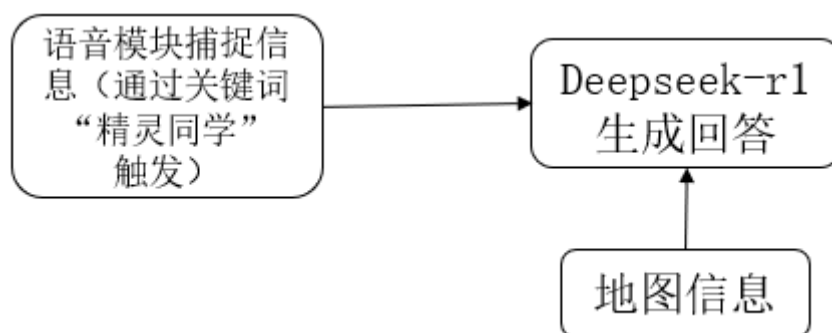


图 13 Deepseek 模块设计图

## 第三部分 完成情况及性能参数

### 3.1 整体介绍

作品整体由开发板主体、摄像头模块、音频输出模块、语音识别模块、LED 模块构，实现基于 RK3588 的大语言模型辅助驾驶系统。

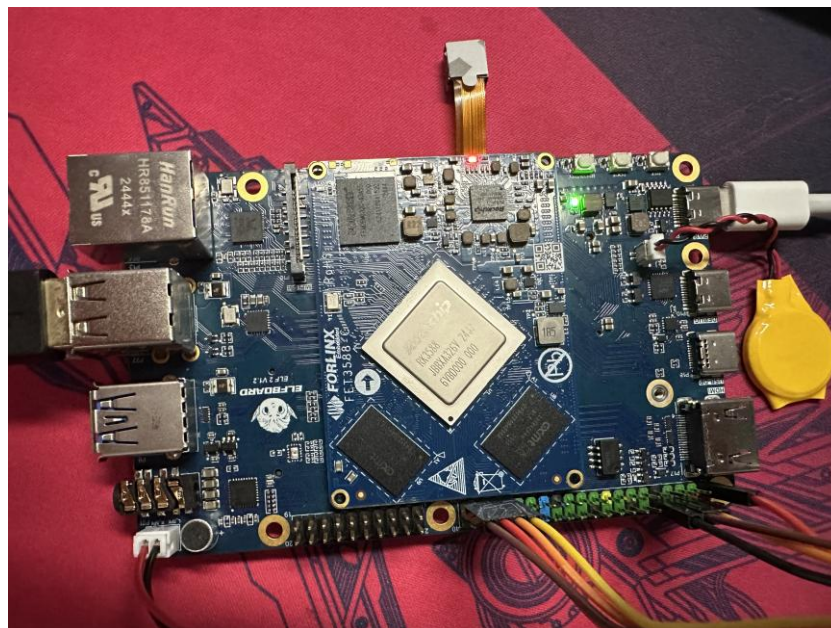


图 14 作品正面图

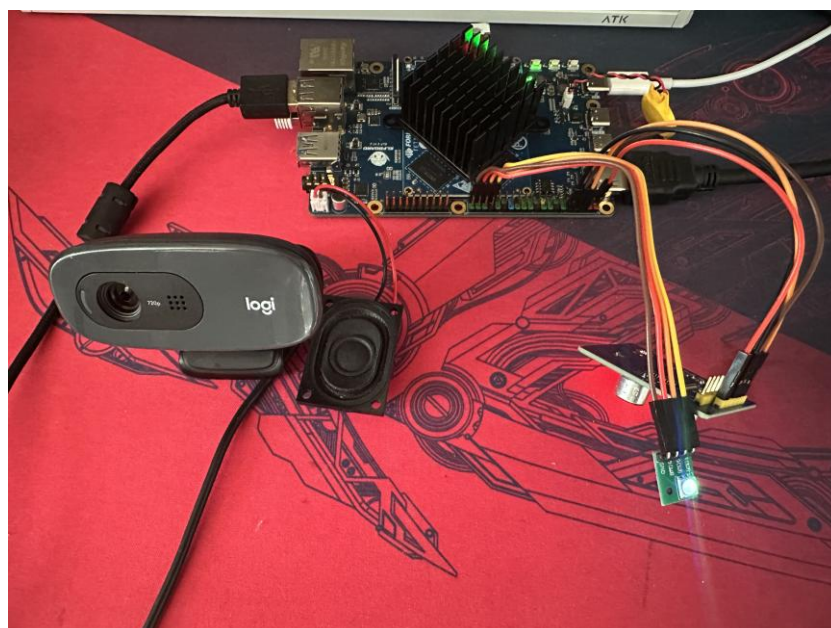


图 15 作品斜 45 度图



## 3.2 工程成果

### 3.2.1 电路成果

LED 四色控制灯代表预测模型给出的四个不同预测结果。

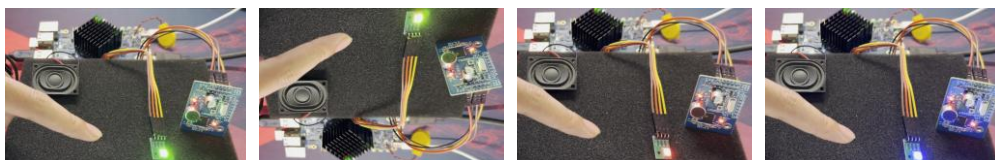


图 16 LED 四色控制灯

语音识别模块上红灯亮代表唤醒成功，紫灯亮表示 Deepseek 正在处理中。

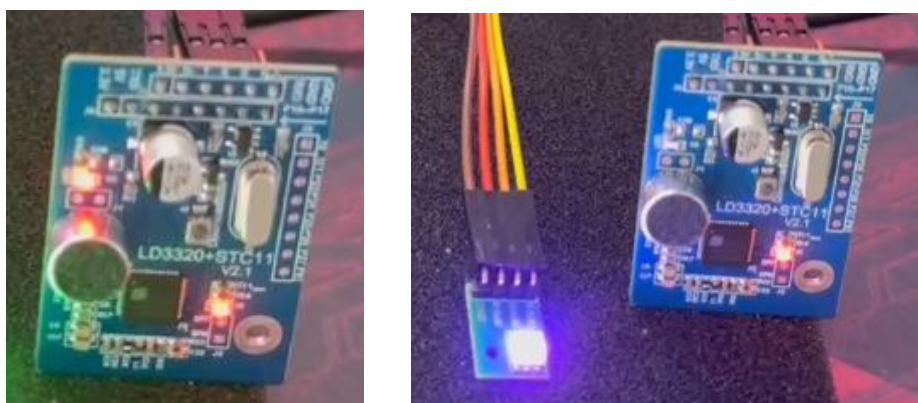


图 17 语音模块

### 3.2.2 软件成果

本作品更倾向应用于车机系统，车机系统主页面可用于音乐播放等其他内容，该页面仅作为系统后台预测演示。



图 18 预测界面图

### 3.3特性成果

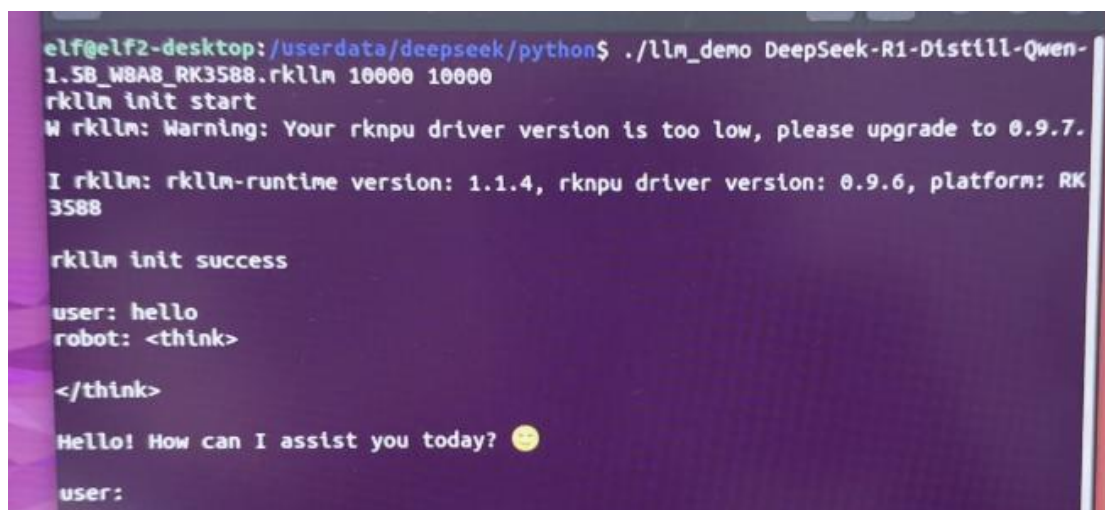


图 19 本地部署 Deepseek 交互

表 5 性能评价表

评判对象	指标	分析
语音识别响应速度	约 0.5 秒	可以满足实时性要求
Deepseek 响应速度	约 40 秒	可以满足实时性要求
道路检测模型推理速度	约 0.2 秒	可以满足实时性要求
道路检测模型识别准确性	准确率约为 90%	集成策略有效降低误判率
道路检测模型资源占用	模型大小约为 32.7MB 显存占用小	适合嵌入式部署
Deepseek 资源占用	内存占用约 2.5GB	适合嵌入式部署



## 第四部分 总结

### 4.1 可扩展之处

多样化外设连接：如多屏异显通过 HDMI 输出实现分屏导航。语音交互等功能，HUD 通过视频接口输出驾驶信息，减少疲劳驾驶，添加红外或夜视摄像头辅助夜间驾驶。

（1）特种车辆专业化：如盲区全景检测与预警、极端环境感知、智能决策与自主作业、人机交互与驾驶员状态监控、集成平台建议等功能。

（2）个性化设计：个性化语音包，个性化车载灯光，实现辅助驾驶后排娱乐两不误等功能的扩展。

（3）专业任务车辆信息提示：通过改变数据库内容，为专业任务车辆人员提供工作信息。

（4）增加对道路信息的处理：本次由于板卡资源限制，未能实现对道路路牌、信号灯等特征的采集处理；可以增强板卡性能来实现对道路其余信息的处理，与语音信息一起传入 Deepseek 实现更为实时的导航提示。

### 4.2 心得体会

在这段时间的研发过程中，我们由实际驾驶中的问题得到灵感去研发一个驾驶辅助系统以解决驾驶中导航系统需要手动输入文档的不便之处，由老师指导后明确了这个系统应该包括哪些模块。在近一个月的制作过程中，组员们各司其职分别完成了软件系统、摄像头模块、喇叭模块、语音识别模块等功能并将其整合为一个驾驶辅助系统。在研发过程中我们也遇到了许多问题，如文字解码有误、因为超出内存无法调用 Deepseek 等，经过讨论与查阅相关资料最后都一一解决了。通过这个研发我们深刻的了解了嵌入式开发板的基本开发流程，学会将模型部署进开发板中，知道怎么调用所接入的外设等收获颇丰。也希望在未来我们也可以设计与研发更多有意思且有用的作品。

## 第五部分 参考文献

- [1] Szegedy C , Ioffe S , Vanhoucke V ,et al.Inception-v4, inception-ResNet and the impact of residual connections on learning[C]//Proceeding of the Thirty-First, AAAI Conference on Artificial Intelligence: Twenth-Ninth Innovative Applications of Artificial Intelligence Conference and Seventh Symposium on Educational Advances in Artificial Intelligence: San Francisco, California, USA, February 4-9, 2017, Volume Five.2017.
- [2] Zhu W , Lan C , Xing J ,et al.Co-occurrence Feature Learning for Skeleton based Action Recognition using Regularized Deep LSTM Networks[J].AAAI Press, 2016.DOI:10.48550/arXiv.1603.07772.Xu K , Ba J , Kiros R ,et al.Show, Attend and Tell: Neural Image Caption Generation with Visual Attention[C]//arXiv.arXiv, 2015.
- [3] Chattha H A .OLD MYTH,NEW LEGEND Deepseek's AI model,rivaling ChatGPT at lower cost and speed,highlights Chinese innovation and empowers the Global South[J].China Report ASEAN, 2025, 10(2):78-78.
- [4] 王曰海.深度学习嵌入式应用开发:基于 RK3399Pro 和 RK3588[M].机械工业出版社,2022.
- [5] 熊洋.基于小词汇量语音识别算法与 TTS 系统的车联网移动终端研究与实现 [D].北京交通大学,2013.DOI:10.7666/d.Y2603188.
- [6] Afridi A H, Yasar A U H, Sheltami T R. Serendipity and LLM-based Recommender System for Smart Transportation[J]. Transportation Research Procedia, 2025, 84: 251-258.
- [7] Liu T, Yang J, Yin Y. Toward llm-agent-based modeling of transportation systems: A conceptual framework[J]. Artificial Intelligence for Transportation, 2025, 1: 100001.
- [8] Wang T, Fan J, Zheng P. An LLM-based vision and language cobot navigation approach for Human-centric Smart Manufacturing[J]. Journal of Manufacturing Systems, 2024, 75: 299-305.