# 基于 ESP32 开发板的语音识别AI小车

## 实验原理：

1. 文本转语音（TTS，Text-to-Speech）：

核心原理是通过神经网络模型将输入的文本转换为自然流畅的语音信号。

1. 数据增强（Data Augmentation）：

通过对现有数据进行变换或扩展来增加数据量和多样性的技术，旨在提高机器学习模型的泛化能力和鲁棒性。

1. **语音KWS（Keyword Spotting，关键词检测）：**

**基于机器学习模型，将语音信号预处理后通过模型训练，实现实时推理关键词检测。**

1. 蓝牙通信（Bluetooth Communication）：

一种短距离无线通信技术，工作在2.4 GHz频段，通过跳频技术和GFSK调制实现抗干扰和高效传输。其核心包括设备发现、配对、连接建立和数据传输，采用分层协议栈支持多种应用场景。

1. ESP32开发板:

通过接收和处理控制指令，驱动电机模块实现小车的运动控制。

## 实验目的：

1. 掌握语音识别技术的基本原理与应用。
2. 学习嵌入式系统开发与蓝牙通信技术。
3. 提升项目实践能力与团队协作能力。

## 硬件设备：

1. 电脑 （运行语音识别模型、发送控制指令）
2. 手机 （作为麦克风，采集语音指令）
3. ESP32开发板 （接收蓝牙指令、控制小车运动）
4. 电机驱动模块 （驱动小车电机）
5. 电机马达 （驱动小车运动）
6. 小车底盘 四轮底盘（承载电机、电池等组件）
7. 电池 （为小车供电）
8. 杜邦线 （连接各模块）
9. 面包板 （方便电路搭建）
10. 数据线 （USB转Micro USB 1根 为ESP32开发板供电和下载程序）

## 软件工具：

1. ChatTTS （用于生成语音数据集）
2. Edge Impulse （用于训练语音识别模型）
3. Arduino IDE （用于ESP32开发）

## 实验方案：

1. 数据采集与预处理语音采集：

通过ChatTTs录制“前进”“后退”“左转”“右转”“掉头”五类语音指令，每个类别数百个wav文件。

数据增强：使用数据增强技术（如噪声添加、变速、变调等）扩展数据集，提升模型的泛化能力，最终获得1800个语音数据。

1. 模型训练：

平台选择：使用Edge Impulse进行模型训练，该平台支持端到端的机器学习流程。

特征提取：从语音数据中提取MFCC等特征，作为模型输入。

模型训练：基于提取的特征，训练语音识别模型，如卷积神经网络（CNN）或循环神经网络（RNN）。

1. 语音识别与指令生成语音输入：

在电脑上输入语音，模型进行识别并输出对应指令（如“前进”“后退”等）。

1. 蓝牙通信：

指令传输：通过蓝牙将识别出的指令发送给ESP32芯片。

1. 小车控制：

电机驱动：ESP32接收指令后，驱动小车电机模块，执行相应动作（如前进、后退、左转、右转、掉头）。

## 实验过程：

1. 数据采集与预处理

语音数据采集：

使用ChatTTS平台生成“前进”、“后退”、“左转”、“右转”、“掉头”五类语音指令，每种指令录制100个样本，通过改变音色、噪声、音量等特征增强语音样本的多样性。最终得到500个WAV文件。

问题：

1）Chat TTS平台只能逐个手动生成语音文件，且只能逐个保存，导致采集工作量大。

2）后续过程中500个数据训练的模型识别效果欠佳。

解决：

引入数据增强技术，通过对各个语音数据进行音调变化、时间拉伸、添加噪声等方法增强，从而扩大数据样本，提高数据的泛化性。

1. 模型训练

模型训练： 将增强后的数据集上传至Edge Impulse平台，选择合适的神经网络模型（如MFCC + CNN）进行训练，并调整超参数以获得最佳识别效果。

1. 模型部署与测试

A.模型部署： 将训练好的语音识别模型部署到电脑端，并编写程序实现语音指令的实时识别。

B.模型测试： 使用测试集对模型进行测试，评估模型的识别准确率、召回率等指标，并根据测试结果对模型进行优化。经测验，模型对测试集的识别效果精准，高达98％以上。

1. 硬件连接与程序编写

A.硬件连接： 按照电路图连接ESP32开发板、电机驱动模块、直流减速电机、蓝牙模块等硬件设备。

B.程序编写： 使用Arduino IDE编写ESP32控制程序，实现蓝牙通信、电机控制等功能。

5.系统联调与功能测试

A.系统联调： 将电脑端语音识别程序与ESP32控制程序进行联调，确保语音指令能够准确识别并通过蓝牙传输至ESP32。

问题：

1. 由Arduino上传至ESP32的烧录过程在初始测验阶段成功，在后续实测中失败。
2. 蓝牙连接ESP32不稳定。

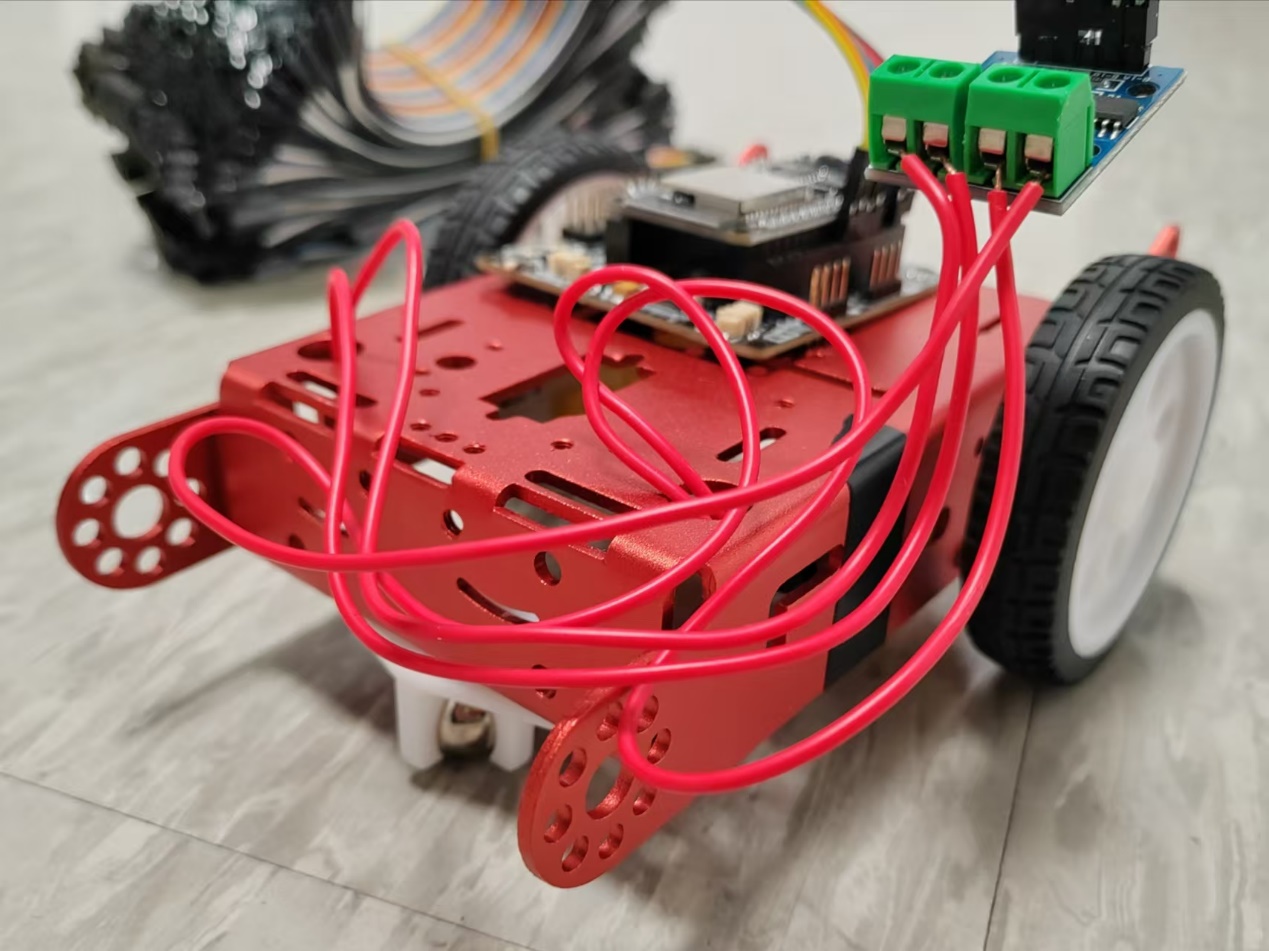
解决：

排除代码错误，推测硬件故障，经逐一检测，最后更换新的ESP32后烧录成功。

B.功能测试： 对小车进行功能测试，验证语音控制小车的各项功能（前进、后退、左转、右转、掉头）是否正常。

## 效果演示：

### 1.小车模型（图片）：



### 2.功能测试（视频）：



## 数据分析：

观察视频测试效果并结合实际多次测验，发现：

1. “前进”“后退”“左转”“掉头”识别效果优秀，“右转”识别欠佳。

2. 由识别到运动时差约为6秒，反应速度有待提高。

3. 小车运动遵循指令，运动轨迹正常。

## 实验反思：

1. “右转”效果欠佳，可能是因为“前进”“后退”“掉头”通过数据增强后样本量充足，而“左转”“右转”数据量相对较少。可以通过数据增强扩大这两类数据样本解决。
2. 缩小识别到运动时差，可以通过
3. 优化识别模型，提高识别速度。
4. 优化蓝牙通信代码，提高传输速度。
5. 将模型载入ESP32，并利用其麦克风，省去传输环节。
6. 优化小车运动代码，提高反应速度。
7. ESP32开发板出现问题可能是运作频繁或测试电压过高，实验过程应该严格控制电压大小，以免损坏芯片。

## 实验收获：

1. 理论与实践结合： 深入理解了语音识别、嵌入式系统和蓝牙通信等技术原理，并成功应用于小车控制，实现了理论与实践的结合。
2. 技能全面提升： 提升了数据处理、模型训练、硬件调试、系统集成等实践技能，同时锻炼了团队合作和项目管理能力。
3. 激发创新探索： 实验过程激发了创新意识和探索精神，为未来深入研究人工智能、物联网等领域奠定了基础。

## 日常应用：

1. 智能安防系统： 结合语音识别和图像识别技术，实现智能门锁、智能监控等安防设备的语音控制，例如通过语音指令开门、查看监控画面等，提高家庭安全性。

2. 智能服务机器人： 在餐厅、酒店、医院等场景，语音识别技术可以用于服务机器人，实现语音点餐、语音导航、语音咨询等功能，提升服务效率和质量。

3. 智能医疗： 语音识别技术可以用于医疗领域，例如语音电子病历、语音辅助诊断等，提高医疗效率和质量。