

请参照下列提纲填报

（一）申请课题信息

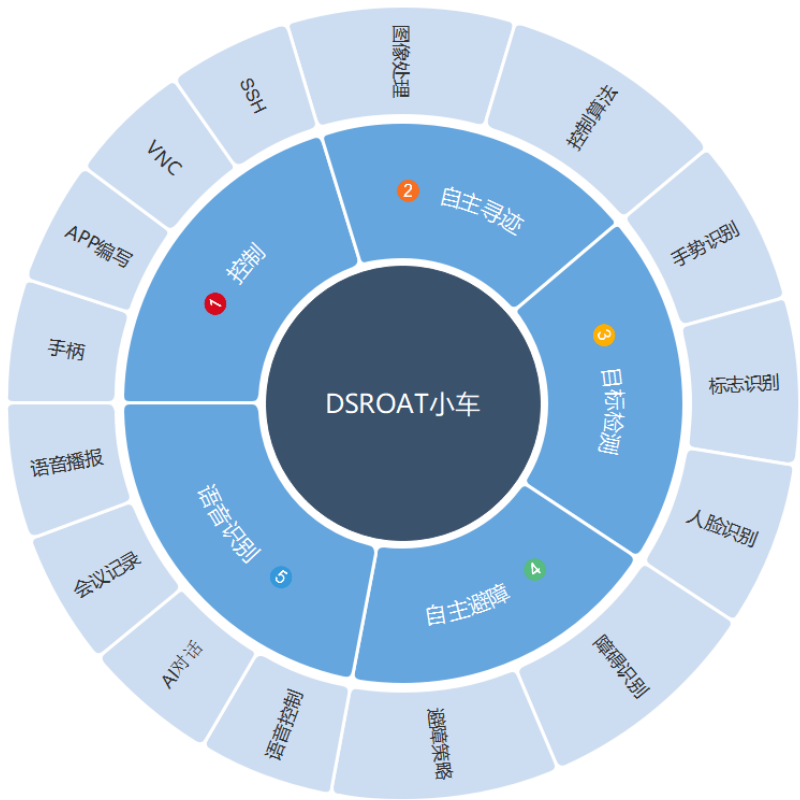
课题名称	基于树莓派的 DSROAT 智能车的研发
预计完成时间	2021 年 5 月

（二）国内外研究状况综述

1. 本项目涉及的关键技术点是什么

近年来，随着硬件系统的不断开发和算法逻辑的深入研究，智能车正出现在市场上，并逐步进入人们的日常生活。智能小车的研发并投入市场使用会对人类的生产与生活产生深远的影响^[1]。

本项目旨在以树莓派主板为开发平台，设计一款满足 DSROAT 四项功能的智能车，主要完成的功能是目标检测(object Detection)、语音识别(automatic Speech Recognition)、自动避障(Obstacle Avoidance)、自动循迹(automotive Tracking)。



本项目设计的智能车主要由电机驱动模块、摄像头模块、车辆控制模块等多个模块在中央控制器芯片树莓派的控制下，完成小车的循迹、避障、执行指令等任务。

本项目涉及的关键技术点主要如下：

- (1) **自动循迹**：使用摄像头模块进行视觉采集，对采集到的图像进行分析，

利用数字图像处理的方法处理图像以获取小车赛道边线，并使用 Stanley、Pure Pursuit、PID 等算法对小车进行控制，以完成自动循线任务。

(2) **自动避障**：需要利用摄像头模块进行视觉采集，采集明显障碍物图像信息，利用模版匹配、神经网络识别等方法进行判断

(3) **目标检测**：分为人脸识别、手势识别、标志识别三个基本任务。需要结合计算机视觉（Computer Vision）的知识与方法，建立应用于目标检测的深度神经网络模型，在现有深度学习框架 Tensorflow 和 Keras 下进行训练，并获得模型训练后的结果。将训练后的模型应用于实际小车中，分别对人脸、手势、标志进行识别，并完成相应任务。

(4) **语音识别**：需要在小车上配备麦克风等外设，使用语音识别技术（ASR）来接收、提取并识别语音，利用一些开源 API，结合自然语言处理技术（NLP）来分析其中的信息，转化为树莓派的指令，完成相应任务。

2. 各个技术点研究的历史与现状

本项目主要涉及了四个技术点：自动循迹、自动避障、目标检测和语音识别，接下来分别介绍四个部分的研究历史与现状。

(1) 自动循迹

智能车的循迹控制是智能车自动驾驶的关键性技术^[2]。一定程度上来讲，智能车的自动寻迹系统可以看作无人车对真实道路的白色指引线的识别与跟踪，属于自动驾驶领域的入门级项目。

国内外自动循迹智能车的发展历程大致可以分为以下三个阶段：阶段一，1954 年美国 Barrett Electronic 公司研发出世界上第一台自动引导车系统，适用于固定路线上的拖车式运货平台。阶段二：二十世纪八九十年代，美国成立了国家自动高速公路系统，研究智能车辆自动循轨的可行性。同时亚洲日本研制自动化车辆导航系统，成立先进巡航协会。阶段三，由美国卡耐基·梅隆大学为主的机器人研究所完成对车辆自主循线的研究，取得显著成就。

本项目使用的方法是使用摄像头模块获取地面图像，利用数字图像处理的方法获取路面信息，判断路线轨迹，而后树莓派控制循迹模块，使用合适的控制器来对小车进行控制，以达成循迹的功能。

(2) 自动避障

智能车的自动避障技术，顾名思义就是智能小车自主躲避障碍物的智能技术。智能车根据障碍物信息，自主避障。

目前智能车自动避障技术可分为两个阶段：一是感知阶段，通过特定模块感知障碍物。现有最常见的感知模块是红外激光模块和超声波模块，具有精度高、速度快的特点。超声波测距技术成熟，成本低，但缺点在于作用距离近，且对反射面有一定的要求。红外激光测距使用的是飞行时间传感器，传感器发射红外射线根据反射信号和原信号的相位差来计算信号的飞行时间，即可换算出距离。但仅利用超声波和红外模块，智能车只能判断前方是否存在障碍物，而不能够判断是哪一种类型的障碍物，因此我们在此次项目中优先考虑利用双目视觉法，运用人眼估计距离的原理，通过摄像头模块获取障碍物的具体坐标以及障碍物类型，获取障碍物的深度图像，进而感知障碍物，为更好地避障做准备。

二是避障阶段。当智能车感知前侧方存在障碍物后，需要设计策略绕过该障碍物，根据视觉处理模块计算出的障碍物的具体位置，选用合适的预瞄点偏移或

开环避障算法完成避障阶段。

(3) 目标检测

传统的目标检测算法一般分为三个阶段：一是在给定的图像上选择一些候选的区域；二是对这些区域提取特征；三是使用训练的分类器进行分类^[4]。候选区域的选择和人工特征因子的制作会很大程度上影响分类器的精度，导致目标检测的精度很低。然而近年来，随着深度学习的广泛应用，基于深度学习的目标检测方法取得了巨大的进步。

基于深度学习的目标检测算法分为两个步骤：目标分类和目标定位。众所周知，目标检测算法已存在一定的历史，传统的目标检测算法使用人工特征提取方法获取图像特征，再通过分类器来进行目标物体的识别与分类。然而，随着深度学习的发展，目标检测进入了一个新阶段。基于深度学习的目标检测算法精度远胜于传统算法，速度也优于传统算法。

一般来说，目标检测器可以根据检测的步骤分为以下两类：第一类是两阶段检测框架（two-stages），另一类是单阶段检测器（one-stage）。检测器还可以根据检测过程是否使用预处理的目标框分为以下两类：一是基于锚框的检测（anchor based），另一类是不基于锚框的检测（anchor free）。最早提出的两阶段基于锚框的检测模型是 R-CNN，首先需要给出目标的候选区域，在每个候选区域内寻找目标。这种方法的优点是精度很高，错误率很低，但缺点很明显，速度很慢，无法应用于实时检测系统中。于是单阶段检测器应之而来。最常见基于锚框的单阶段检测器有 YOLO 和 SSD 系列，其中 YOLO 系列以其高速、准确的特点闻名于目标检测界。众所周知，单阶段检测器的速度快于两阶段，但精度明显低于两阶段模型。而 YOLO 采用的方法不同于 R-CNN，不需要提出候选框的步骤，而是直接一步到位全局回归出目标的位置和种类，大大提高了速度，但同时又保持了较高的精度和召回率，可谓是在二者之间实现了平衡。但 YOLO 系列仍是基于锚框的，锚框大小是一个重要的超参数，也就是说需要人工事先设置锚框（预选框）的大小，才能使模型更快收敛。倘若锚框的设置不合理，会导致最终检测效果的崩塌。因此，我们常常使用聚类分析法，将训练集中的所有目标进行聚类，得出各类的锚框大小，再输入至模型中，以保证检测效果。因此，人为设定的锚框大小极其重要。

为克服该超参数的影响，近两三年研究者转向研发无锚框的目标检测。截止日期，最常见的无锚框单阶段的目标检测方法有 CornerNet、CenterNet、FCOS、TTFNet 等。其中，TTFNet 当属速度最快、精度最高的检测器。该检测器不需要锚框，而是直接回归目标点，以及以目标点为中心的左右上下的宽度来完成目标检测。无锚框的目标检测是未来研发的主要方向。

由于无锚框的目标检测技术尚未成熟，为更好地完成我们的项目，暂定选择采用最新版 YOLOv5 来完成目标检测，考虑树莓派上的性能限制，我们也将传统方法与分类器结合的方式作为备选方案。

(4) 语音识别

语音识别的目标是将人类的语音中的词汇内容转化为计算机可读的输入。语音识别的发展过程经历了 3 个阶段^[5]，一是孤立词识别，二是连接词识别，三是大词汇量下连续语音识别。

自上世纪 50 年代开始，着手于最简单的数字识别任务，语音识别领域进入

研究者的视野。到 80 年代，研究者们采用统计分析的方法使连续语音识别成为可能。在我国，50 年代末有研究者采用电子管电路，对英语中的元音进行尝试识别。90 年代，清华大学和中科院自动化所等单位在汉语听写机原理样机的研制方面取得有效成果^[5]。

在国外，从 1980 年开始，CMU 使用 GMM-HMM 隐马尔可夫模型实现非特定人连续语音识别。GMM-HMM 模型分为两步，首先建立语音字典，建立语音和音素状态之间的多对一关系，建立 GMM 模型进行拟合。再对 HMM 模型进行估计，得到模型参数。该模型取得了较好的结果。然而在 80 年代后期，人工神经网络被引入语音识别领域，于是多种新方法诞生，深度神经网络 DNN、卷积神经网络 CNN、循环神经网络 RNN 都对语音识别做了一定的贡献。近年来，研究者发现引入注意力机制的端到端训练能够使语音识别模型的训练直接从语音转录为文本，省去了声学模型和语言模型，大大提高了识别效率。

3. 本项目主要存在的问题和难点

本项目涉及多个领域，多门专业知识，因此难度系数不低。我们将分为四个部分进行问题和难点的叙述。

（1）自动循迹部分

①相机标定。

本次项目利用摄像头视觉获得场景特征，以达到循迹目的。然而，摄像头的类型并不清楚，因此需要进行相机标定。相机标定需要采集大量图片，利用图像处理算法得到摄像头响应的内参和畸变系数，并保证误差范围，这是本项目的一个难点。同时，使用的标定方法和图像处理算法得到的结果会极大的影响后续任务的进行，这是项目的一大问题。

②控制器的选择与参数整定

在获得赛道特征后，需要对小车进行控制。如何选择控制器，使控制小车在赛道线中心位置运动，不发生偏离现象，特别是在左转和右转时，对速度的精确控制是一大难点。

（2）自动避障与目标检测

①数据标定

要完成目标障碍物检测，大量的数据标定是必不可少的。图像的标定是一个耗时耗力的工程，考虑到项目的时间紧迫性，这是一个问题。

②传统算法和深度学习的抉择

就算法性能相比，传统算法远不如深度学习算法，深度学习带来的高准确度和高召回率多方面碾压传统算法。但考虑到树莓派小车硬件配置，传统算法的优点就尤为突出：速度快。如果采用深度学习方法进行目标检测，检测的速度会大打折扣，可能无法满足小车的实时检测效果或存在延迟效应。如何在传统算法和基于深度学习的目标检测算法之中进行权衡是本次项目的一个问题。

（3）语音识别

缺失麦克风、USB 声卡等，我们尽快计划在淘宝购买外设使用。

参考文献

[1] 张秉森, 马吉忠, 杨一飞, 张敏, 毛汉奎, 吴丽丽. 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(03): 23-24.

[2] 吴畏, 赵华东, 张耀辉. 智能驾驶车辆自动循迹的研究[J]. 汽车实用技术, 2020, 45(19): 35-37.

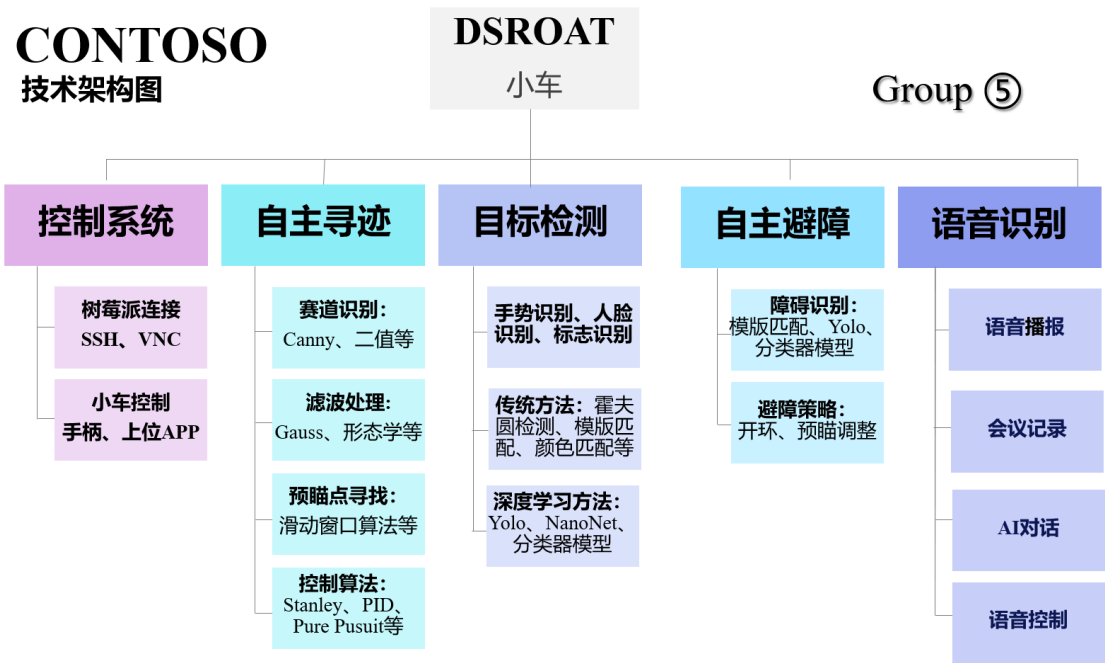
[3] 周伟, 李军. 自动驾驶车辆避障路径规划研究综述[J]. 汽车工程师, 2018(05): 55-58.

[4] 陆峰, 刘华海, 黄长缨, 杨艳, 谢禹, 刘财喜. 基于深度学习的目标检测技术综述[J]. 计算机系统应用, 2021, 30(03): 1-13.

[5] 鱼昆, 张绍阳, 侯佳正, 张少博. 语音识别及端到端技术现状及展望[J]. 计算机系统应用, 2021, 30(03): 14-23.

(三) 拟采取的技术途径和关键点

1. 本项目实施中的技术架构与关键点



2. 创新点和研究重点

本次实验的创新点在于，使用简易的树莓派小车模拟真实的应用环境，并以demo 的模式进行封装与模拟。

在本次实验中，我们可以模拟无人车在真实道路上的寻迹巡线与对标识牌的识别、模拟 AGV 系统在工厂中的目标检测与到点卸货、模拟服务机器人在实际生活中的跟人送货、模拟酒店无接触机器人的手势识别与声音控制、模拟家庭生活中的语言播报与远程控制等等，在当今机器人与人工智能被广泛推崇与应用的时代，我们使用简单的树莓派小车对这些实际场景进行了模拟，以到达相应的应用层面。

（四）研究计划和分工

1. 研究计划

为克服项目难题，我们小组设定了如下的研究计划。



(1) 2021年3月8日至3月14日：小组各位成员对四个技术点（自动循迹、自动避障、目标检测、语音识别）分别进行文献调研，查看业界现有的思想与算法，并进行汇总，为后续开发打下坚实的基础。

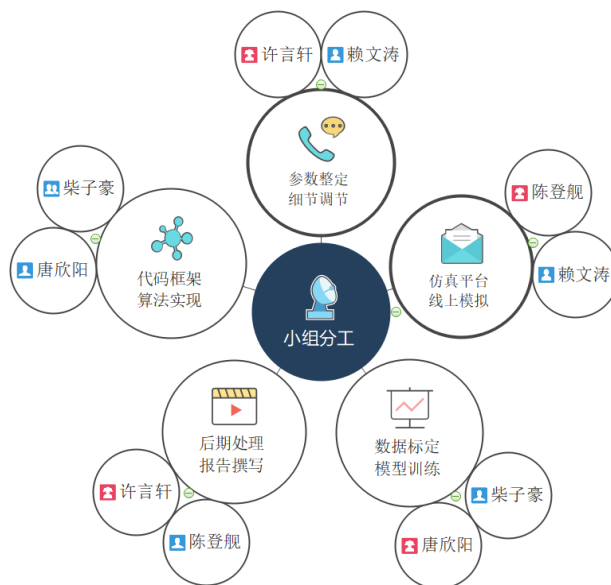
(2) 2021年3月15日至3月21日：小组成员组织多次线下集中讨论，共同参与研发，讨论需要使用的算法，制定合理步骤，完成开题报告，并熟悉树莓派相关平台。

(3) 2021年3月22日至5月23日：小组成员共同参与开发，分别深入研究四个技术难题，各个击破，进行多方面调试，初步完成整个项目内容。

(4) 2021年5月24日至5月30日：小组成员对项目的研发结果进行分析，讨论优点和不足之处并进行合理完善，完成结题报告的撰写，准备答辩。

2. 小组分工

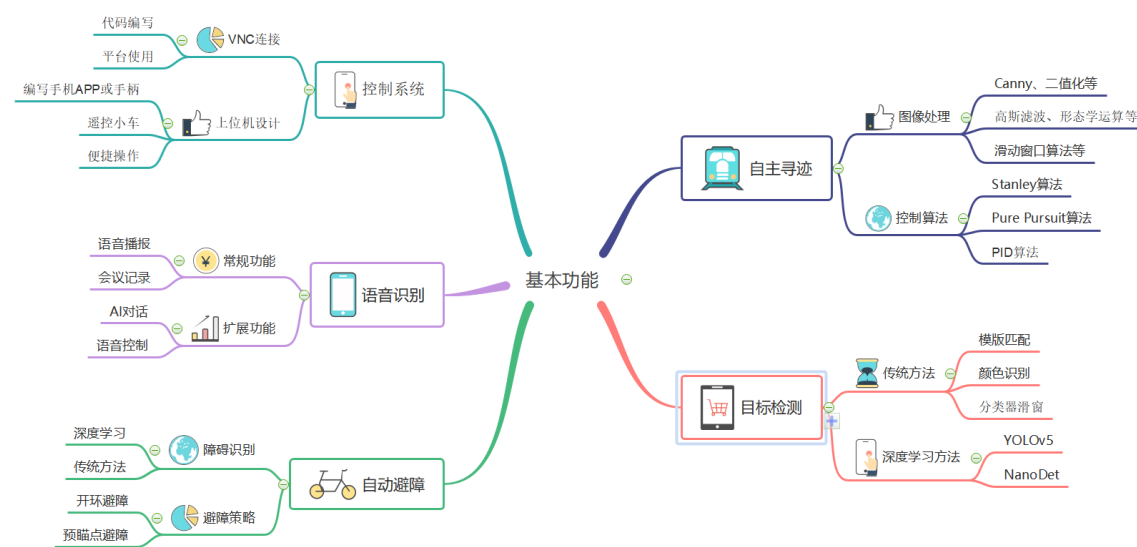
本次项目拟分工情况如下：



- (1) 实际应用调研：5 人
- (2) 算法设计与代码实现：2 至 3 人
- (3) 线上仿真平台：2 至 3 人
- (4) 数据采集与标定：2 至 3 人
- (5) 周报与报告撰写：1 至 2 人

(五) 预期成果

以树莓派主板为开发平台，设计完成满足 DSROAT 四项功能的智能车，以模拟真实的应用环境为主要目的，完成目标检测、语音识别、自动避障、自动循迹等基本功能。



自动寻迹功能模拟真实无人车在高速公路行驶时的无人驾驶或者说道路保持，利用车载的摄像头模块对实验场地进行赛道特征的识别，使用 Stanley、Pure Pursuit、PID 等算法对小车进行控制，保证速度与精度地完成本次自动循线任务。

目标检测功能完成人脸识别、标志检测、手势识别等基本功能。人脸识别旨在在设定目标后，模拟真实服务机器人，完成对目标服务人员的识别跟踪。手势识别目的是模拟无接触条件下对机器人的控制，例如，酒店送餐机器人在特定的手势下完成开关灯、左转右转、取餐等动作。标志检测是为了实现实验场地中，也是为了模拟真实 AGV 环境中的实际应用，在识别到左转、右转、限速等标志时进行相应的操作，以使得控制系统更加完备。

自动避障功能实现在识别到特定标志后，执行相应的到站转弯与停车操作，旨在模拟真实 AGV 环境到到站功能。

语音识别部分，计划完成如下功能：语音播报，如“左转”、“停车”、天气播报等命令；会议记录，在会议过程中快速记录文字；AI 聊天，模拟流行的聊天机器人，如图灵机器人、小爱、Sir 等完成基本的对话功能，查阅资料并播报；语音控制，如使用命令“关灯”、“前进”、“左转”等，达到对树莓派及扩展系统的相应控制。