

# 基于深度强化学习的无人驾驶决策算法研究

宋北丞，韦洁

南方科技大学

**摘 要** 行为决策系统能够综合环境及自车信息，使自动驾驶车辆产生安全合理的驾驶行为，是实现无人驾驶的核心。强化学习算法采用一种自监督学习的方式，使自动驾驶车辆的决策系统在与环境的交互过程中，通过不断改进自身策略自主学习到最优的决策模型，为构建有效的决策系统提供了方向。文中介绍了目前无人驾驶决策算法的基本框架与一些基于学习的与自动驾驶相关的算法与一些传统的决策算法，并介绍了近年来基于深度强化学习的决策相关算法的研究进展。文中也列举了决策算法在无人驾驶汽车中的具体应用，从礼让行人、紧急避险、变道以及通过红绿灯路口四个方面详细说明了决策算法的具体决策过程。文章最后从尚可提升之处与当今时代有利于无人驾驶发展的有利之处两个方面提出了决策算法的研究前景。

**关键词：**无人驾驶；决策算法；深度学习

## Research on Autonomous Driving Decision Algorithm Based on Reinforcement Learning

Beicheng Song and Jie Wei

Southern University of Science and Technology

**Abstract** The behavior decision system can integrate the environment and self-vehicle information to make the self-driving vehicle produce safe and reasonable driving behavior, which is the core of the realization of driverless driving. Reinforcement learning algorithm adopts a self-supervised learning method, so that the decision-making system of autonomous driving vehicle can learn the optimal decision-making model through continuous improvement of its own strategies in the process of interacting with the environment, which provides the direction for the construction of an effective decision-making system. This paper introduces the basic framework of autonomous driving decision algorithms, some learning-related algorithms and some traditional decision algorithms, and also introduces the research progress of decision related algorithms based on deep reinforcement learning in recent years. The paper also lists the specific application of the decision algorithm in the driverless car, and explains the specific decision process of the decision algorithm in detail from four aspects: comity to pedestrians, emergency avoidance, lane change and crossing traffic lights. At last, the paper puts forward the research prospect of decision algorithm from two aspects, which can still be improved and the advantages which are conducive to the development of unmanned driving in the current era.

**Keywords** Autonomous driving; decision algorithm; reinforcement learning

随着科技的发展,自动驾驶因其在军事、商业以及民用领域广泛的应用前景与巨大发展潜力而成为近年来的一大研究热点,而无人驾驶汽车则是该项技术蓬勃发展背景下的一项新兴产物。无人驾驶汽车是一个集外界环境感知与信息收集、行驶路径规划与决策制定、驾驶动作执行与运动控制为一体的一项综合智能系统。环境感知系统可以通过车载摄像头、车载雷达等设备对外界环境信息进行收集,并利用计算机视觉与信息融合技术对收集到的信息进行处理。决策系统能够通过对处理后的外界信息以及车辆自身的状态信息进行判断并做出合理的驾驶行为决策。运动控制系统可以根据决策系统做出的决策信号控制车辆的运动。决策系统作为整套无人驾驶系统的核心与“大脑”,是自动驾驶最为核心的技术,对于整个系统的性能与效率有着举足轻重的地位。因此决策系统,特别是安全高效且智能的决策系统成为自动驾驶技术领域的一大研究热点。

强化学习是一种新兴的人工智能算法,智能体在与环境的不断交互反应中得到不同的反馈信息,并基于这些反馈信息不断改善自身的算法策略以及部分参数,从而在不断的学习与调整中得到能够适应环境达到目的的最优算法,强化学习由于其显著优越性与高度灵活性在很多领域中都展现出了较强的应用潜力。深度学习类似于传统的机器学习,但是深度学习可以利用神经网络结构学习数据特征并进行回归或分类,深度学习在图像处理方面有较强的优势。将强化学习与深度学习相结合并应用于自动驾驶的决策算法中有助于构建一个具有高自由度的车辆行驶决策模型,通过多次模拟各种真实环境下的车辆行驶场景来训练并优化该决策模型,最终可以得到一个能够高效做出最优决策并对突发情况做出有效应对的自动驾驶决策模型。

目前已有许多在相关领域的算法研究综述,分别专注于算法模型的不同方面进行归纳总结与展望,在文献[1]从决策精度、决策广度、应对不确定因素三个方面对于基于强化学习的决策算法模型研究进展进行了简要介绍;文献[3]简要介绍了基于学习的自动驾驶决策算法的基本概况、目前应用以及未来研究前景;文献[3]则介绍了无人驾驶决策规划领域的新趋势,如端到端学习以及交互式路线规划等。这些综述都基于某些特定方面对于现有的自动驾驶决策算法进行分类,并对每一类算法进行分析与比较,但由于缺少具体的应用场景所以直观性较弱;本文不仅对于自动驾驶的决策算法相关信息进行了详细的阐述,还结合某些特定应用场景更加直观地体现出基于强化学习的决策算法的优势与运行逻辑,并对于现有决策算法在自动驾驶汽车中的应用进行了系统的分析与前景展望。

本文共分为四个部分,在第一部分中将简要介绍目前无人驾驶决策算法的简要概括与基本框架,第二部分介绍了一些基于学习的与自动驾驶决策相关的算法以及一些经典的决策算法,第三部分介绍了近年来基于深度强化学习的决策相关算法的研究进展,第四部分结合一些具体场景总结了决策算法在现有无人驾驶汽车之中的应用,第五部分讨论了自动驾驶汽车决策算法未来的发展方向以及研究重点。

## 1 决策算法的基本概述

1.1 决策系统的基本框架

自动驾驶系统中的决策部分是感知环境与信息收集部分以及驾驶动作执行与运动控制部分的过渡，通过处理感知系统收集到的关于外界的信息从而做出相应的决策并由执行系统进行实现。对于决策系统来说，其输入即为外界环境信息以及车辆自身的状态信息，而输出即为一系列动作执行命令，包括输入运动执行系统的高级与低级行为控制命令。决策系统的基本框架如图 1 所示，输入信息包括由车载传感器收集到的周边的环境信息、车辆自身状态信息以及高清地图信息，而输出的信息包括高级与低级行为控制命令。高级行为控制命令包括超车、变道等，低级行为控制命令包括控制车辆的速度与加速度等。

1.2 决策系统的基本准则与约束

决策系统的目的是通过一系列的标准与规范得出类人的安全高可靠的驾驶决策。而决策系统的基本准则则可以简要归纳为五个方面：合理正确地生成决策，决策的实时性好，行车安全与效率的平衡，车辆行驶稳定性好，高故障检测能力与突发情况应对能力。

为了能够得到更加高效可靠的决策，决策系统需要将多方面因素纳入决策方法的约束中，包括车辆周边的环境信息如道路信息、周围车辆与行人的行为预测、交通标志信息等，当地的交通法规信息，车辆自身状态，每一决策导致的后续结果，历史决策结果参考，驾驶道德等。

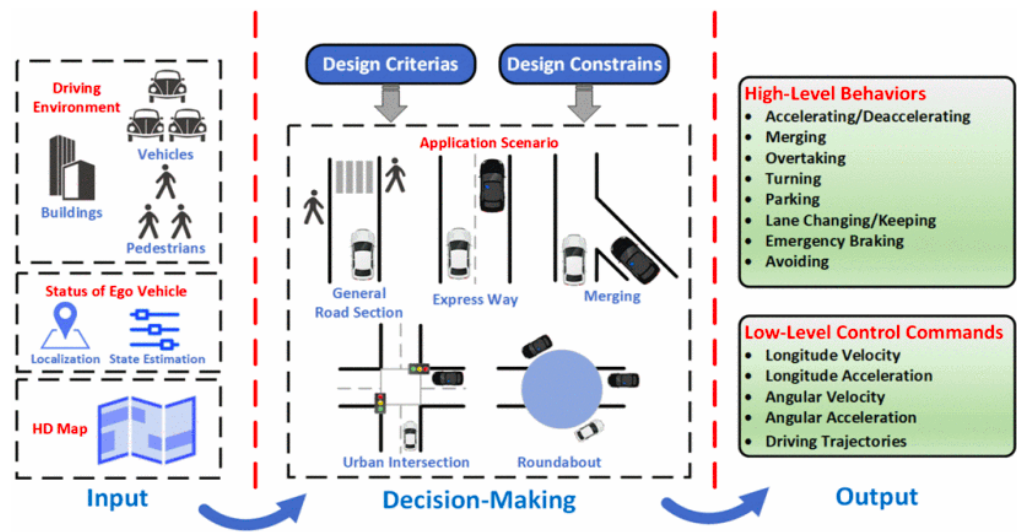


图 1 决策系统基本框架 (Liu,2021)

2 基于学习的决策算法与传统决策算法的简要介绍

2.1 传统决策方法

一般来说，传统的决策方法根据决策依据可以大致分为三类：基于规则的决策方法，优化决策方法，基于概率的决策方法。

2.1.1 基于规则的决策方法

基于规则的决策方法根据由各种交通法规、驾驶知识与经验等构成的规则数据库以及车辆自身状况等信息进行决策。这类决策方法具有较好的决策广度与较强的可积性与适应性，但是由于缺少决策深度所以很难在复杂的驾驶环境下做出正确高效的决策，且对于动态驾驶环境的鲁棒性较低。

此类决策方法中最具代表性的是有限状态机（FSM）。FSM 是一种具有离散输入与输出的数学模型，能够根据对外界环境的响应生成对应的动作，智能体的状态也对应地发生了改变。FSM 方法目前已经运用在一些传统的自动驾驶车辆中。

### 2.1.2 优化决策方法

优化决策方法通常通过奖励函数或者效用函数来生成决策结果。这类方法可以更好的将不同交通组成部分之间的互相作用进行建模，并且可以生成更加优化的决策结果。但是所谓的“优化策略”对于不同的场景往往有不同的含义，因此也会影响系统做出决策。

模型预测控制（MPC）可以生成自我车辆的高级行为，而其他车辆都会收到自我车辆行为的影响。博弈论是另一种解决决策问题的优化决策方法，博弈论首先提出所有智能体都将采用“最优策略”的假设，并将根据其他智能体的相应策略产生行为。

### 2.1.3 基于概率的决策方法

基于概率的决策方法可以根据数学中的概率论来生成每一行为对应的结果，并需要通过建立概率模型来得到行为决策。这类方法很便于与其他类型的决策方法相结合，但是这类方法具有较低的计算效率并且很难在较为复杂的环境下得到最优决策。

## 2.2 基于学习的决策方法

基于学习的决策方法是指通过使用人工智能技术实现自动驾驶决策，通常需要首先建立一定的数据样本，之后采用不同的学习方法或网络结构，根据不同的环境做出合理决策，实现系统的自主学习。基于学习的决策算法主要分为三类：基于统计学习的决策算法，基于深度学习的决策算法，基于强化学习的决策算法。

### 2.2.1 基于统计学习的决策算法

基于统计学习的决策算法可以通过处理大量的训练集数据从而达到类人的决策水平。此类决策算法具有良好的通用性，很适合用于在拥有大量环境数据下的简单情景决策；但是此类算法具有较低的决策准确度，且需要大量的训练数据集来进行自我学习。

基于统计学习的决策算法中较为经典的有 SVM 和 AdaBoost 等。SVM 可以将相对位置和行驶速度作为输入进行变道决策，在考虑变道收益、安全度等约束条件的情况下，结合贝叶斯参数优化能够更好地确定变道模型。该算法最终在实际运用检验中得到了较高的准确率。AdaBoost 可以用来进行车辆插入场景下的风险评估决策，该算法将车辆的速度作为输入，将碰撞距离（DTC）作为输出，能够实现该情况下的安全机动

。

### 2.2.2 基于深度学习的决策算法

基于深度学习的决策算法框架与传统的机器学习算法类似，但是主要区别在于深度学习使用神经网络结构来处理数据提取数据特征，并对数据进行分类或者回归产生对应结果。由于深度学习算法在处理图像是具有优势，在一些特定场景下具有较高的决策准确度；端到端的决策系统保证了对于环境信息的充分利用，被用于自动驾驶汽车的决策系统。该类方法通常将感知系统收集到的数据作为输入，通过神经网络输出低级控制命令。

由于视觉传感器的高性价比，其被广泛配备在自动驾驶汽车上，因此一些端到端的系统就是基于视觉图像进行的。基于卷积神经网络（CNN）进行端到端系统的决策培训，可以成功实现不同环境下是否转向的决策任务。注意力分支网络（ABN）将输入的原始视觉图像进行处理生成注意力图，再将注意力图与原始图像的卷积特征和车辆的预期速度相结合，得出是否转向与是否踩油门的决策命令，实现端到端的决策任务。

### 2.2.3 基于强化学习的决策算法

强化学习是目前最常用的基于学习的决策算法方法之一，强化学习通过尝试各种行为来得到最大获利的行为决策，算法中的参数与行为可以根据奖励函数进行调整。基于强化学习的决策算法更适合对于不确定的或动态的环境进行建模，且灵活的算法框架使整个方法更具扩展性；但是此类方法非常依赖于其内的奖励函数且较为不稳定，在深度强化学习算法下可能会导致过拟合。

基于强化学习的决策算法通常依赖马尔可夫决策过程（MDP）来描述智能体于环境之间的互作状态，同时还提出了部分可观测马尔可夫决策过程（POMDP）来描述更加真实的状态空间。

## 3 近年来基于深度强化学习的决策算法的研究进展

近年来随着深度学习与强化学习领域的快速发展，深度强化学习（DRL）也展现出在智能决策方面的巨大潜力。深度强化学习的核心思想是将神经网络整合到强化学习中，代表性的方法有 Deep Q Network (DQN)、Deep Deterministic Policy Gradient (DDPG) 和 Asynchronous Advantage Actor-Critic (A3C) 等。深度强化学习算法结合了深度学习的表征优势和强化学习的决策优势，显著地提高了决策的规模与维度，通过处理高维信息输入实现了高精度的决策规划。图 2 展示了深度强化学习的基本框架，通常由四部分构成：交互环境 E、智能体 Agent、状态转移的规则、奖励函数及奖励规则。

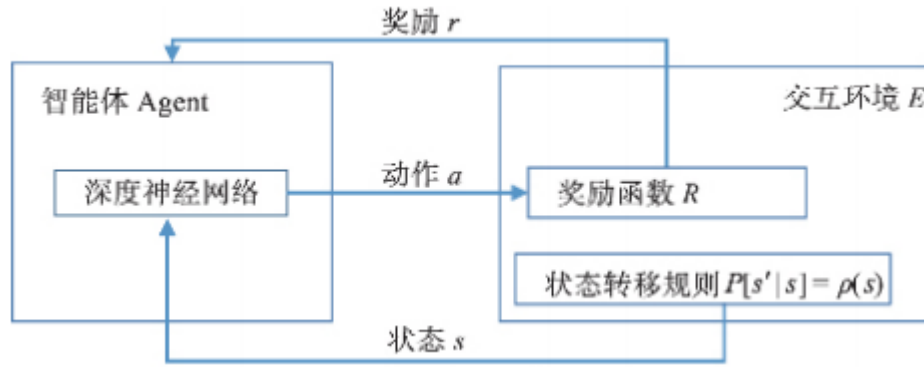


图 2 深度强化学习框架 (Huang, 2020)

### 3.1 DQN 算法

深度强化学习往往将问题建模为部分可观测马尔可夫决策过程 (POMDP)，这是一种离散随机控制过程，在该过程中环境的状态是部分可观测的，并且当前的行为会对未来的状态产生影响，并进一步影响未来的奖励函数输出结果。这意味着智能体在进行决策时需要考虑所做出的决策对于未来产生的影响。而深度 Q 算法 (DQN) 就是通过将卷积神经网络 (CNN) 与 Q-Learning 算法相结合得到的一种基于值的决策算法，智能体可以给每一对状态与动作预测一个对应的 Q 值，并从所有的集合中选取具有最高值的动作作为最优策略，同时智能体还可以利用这些值来训练模型进行优化。DQN 的主要目标就是通过不断学习优化得到最优 Q 函数，通过不断实验对不同状态下的不同行为进行赋值并计算奖励函数输出，并在优化过程中使用贝尔曼方程不断更新神经网络从而得到最优 Q 函数。目前有许多原始 DQN 的变形模型，如 Double DQN 与 Dueling DQN，通过将动作与值预测分流来达到更加高效与稳定的学习效果。

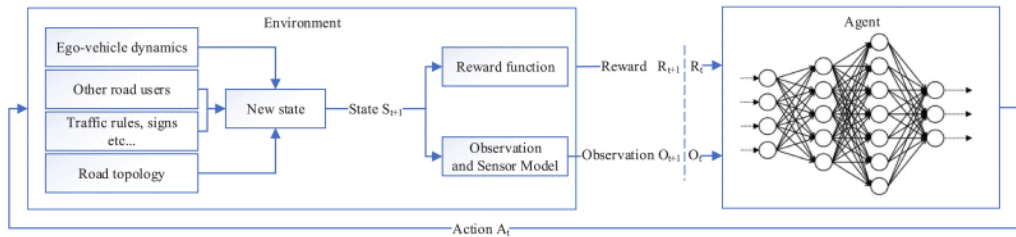


图 3 POMDP 模型 (Aradi, 2020)

### 3.2 DDPG 算法

深度确定性策略梯度算法 (DDPG) 基于 DQN 算法理念进一步提出，能够将深度强化学习拓展到连续的动作域中，实现更好的性能。相较于离散型动作输出的 DQN 模型，连续型输出的 DDPG 决策模型通过端到端的驾驶决策具有更加优越的决策效果。DDPG 算法属于 Model-Free 深度强化学习算法，采用 Actor-Critic (AC) 结构体系：Actor 网络，Critic 网络以及两个网络各自对应的目标网络共同构成体系的网络部分，并且还有随机噪声以用于增加系统的环境探索能力和经验回放池 (Replay-Buffer)，并通过离线策略 (Off-Policy) 的方式对整个网络进行训练。

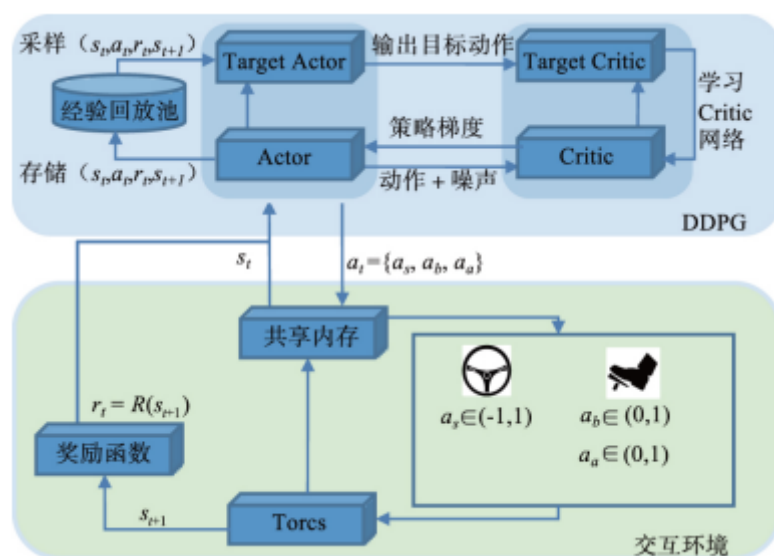


图 4 DDPG 系统模型 (Huang, 2020)

#### 4 决策算法在现有自动驾驶汽车中的应用

决策算法在开始其实并不局限于自动驾驶领域，其决策的能力在很多的方面都有应用，并且在深度学习算法将深度学习的表征能力与强化学习的决策能力相结合之后，显著的强化了学习空间的规模和维度。就比如在阿尔法狗程序的开发上，首次将卷积神经网络和 Q-Learning 算法结合以后提出了深度 Q 网络算法。被训练以后的深度 Q 网络算法已经在某些游戏中的表现高于了人类玩家，在其和蒙特卡洛树搜索算法结合之后，阿尔法狗便被创造了出来，并且在各大比赛的表现中屡次击败人类玩家，如今它的学习水平仍然在不断的进步。更多的在不同领域的决策算法的应用数不胜数，接下来我还是主要介绍一下决策算法在自动驾驶汽车的应用。

##### 4.1 自动驾驶中对于礼让行人的决策：

首先是信息收集:选取不受红绿灯控制的斑马线路段，收集真实车辆驾驶员向行人让行的数据，包括车速、车与行人的距离、行人在斑马线处的位置、车辆减速等。第二步是信息处理。对第一步收集到的数据进行筛选，剔除驾驶员突然刹车行为和行人强制停车行为的数据，保留驾驶员主动让行和减速行为的数据。第三步是建立规则:将第二步中保存的数据进行分类组合，根据分类组合后的数据建立司机向行人让行时的逻辑模糊规则。第四步是规则应用:应用第三步中建立的驾驶员向无人驾驶汽车让行时的逻辑模糊规则，控制无人驾驶汽车在没有红绿灯的斑马线处向行人让行。从这里，我们可以看到自动驾驶仪对行人让行行为

的学习过程。该方法根据无人驾驶汽车当前的速度和与行人的距离，自动实现符合真实驾驶员操作的减速，使车辆在行驶过程中更加稳定和拟人化。

#### 4.1.1 自动紧急刹车系统 AEB:

AEB系统和POMDP规划器一起工作。它使用了POMDP规划器生成的驾驶轨迹来计算碰撞风险。如果碰撞不可避免会发生，AEB系统就会触发，进行紧急刹车，紧急停止具有最高优先级，并能够否决掉POMDP的计划。由于AEB系统能够要求更强的制动干预，因此将POMDP规划器与AEB系统相结合可以获得更高的平均速度。下图显示了紧急制动干预的数量，当行人出现的概率较高时，紧急制动干预的数量就会减少。随着干预次数的减少，系统在通过阻塞区域时变得更加保守。

配备AEB系统的POMDP规划者能够更快地行驶，但需要紧急制动干预以避免碰撞。POMDP计划器在块区域前减速。以低于 50 公里/小时的速度行驶，避免与被阻挡的行人相撞。在这种情况下，AEB系统无法避免以 50 公里/小时的速度发生碰撞。在高速行驶时，反应时间不足以让车辆停下来。障碍物前的速度过高，减速不足，这说明了基本AEB系统的优势。AEB系统不能避免所有碰撞，但两个POMDP规划器可以避免所有碰撞。POMDP规划器和AEB系统的联合实施可以更快地通过障碍物。平均速度更高，但会触发四次紧急制动干预。

#### 4.2 自动驾驶中对于紧急避让的决策:

在自动驾驶避免紧急情况的情况下，需要以下模块进行决策:

环境感知模块:用于获取环境信息，包括:车辆位置参数，障碍物与无人车之间的距离障碍物的形状信息，无人车与其他物体之间的相对距离，相对速度，运动角度和方向，障碍物的二维图像信息，收集车道线，红绿灯信息，无人车制动压力信息，车辆减速，加速信息，车辆质量信息，阀门位置信息、车速信息;根据障碍物的二维图像信息和雷达获得的点云数据，可以推断出物体的三维信息。

危险识别模块:根据获取的环境信息，分析车辆与障碍物之间是否存在碰撞危险，输出碰撞判断信息，确定障碍物类别，输出障碍物类别信息。在判断车辆与障碍物是否会发生碰撞时，会根据获得的环境信息，计算无人驾驶车辆与障碍物之间的安全距离，判断是否发生碰撞事故。如果是，则继续跨越障碍的决策步骤。否则，车辆不需要激活套期保值，车辆按照原计划进行。在判断障碍物类别时，会根据获得的三维信息判断障碍物类别，并判断障碍物是否是易损物体。如果是，则判断障碍物所在车道为不可行的任务路径，输出碰撞判断信息和障碍物类别信息，进入智能决策步骤继续执行。否则，判断障碍物所在车道为可行任务路径，输出碰撞危险信息和障碍物类别信息，进入智能决策步骤继续执行。



智能决策模块:根据获得的环境信息、碰撞判断信息和障碍物类别信息, 构建可行任务池, 分析可行任务事故严重程度, 确定最优可行任务。

越障判断步骤:识别并判断障碍物是否为可跨越障碍;若障碍物为可跨越障碍, 则车辆不需要启动规避危险, 车辆按原计划行驶;如果没有, 请进入确定障碍类别的步骤。

#### 4.3 自动驾驶中对变道的决策:

整个决策过程分为信息采集及信息传输, 模拟计算和最终决策控制四个部分。在信息采集方面, 通过对路面及路边障碍物、行车的车道边界, 当前车辆方圆若干距离内的周边车辆的行驶速度、相对本车的距离等外界的道路信息进行记录, 并且将此进行稍微粗糙的预处理, 提取有效的数据部分传输给主控制模块等待下一步的操作; 接下来执行模拟计算任务的就是主控制模块, 其主要通过对于上述采集的信息与设定的安全信息进行模拟对照并且计算, 得出相应情况下的各个关键数值后, 与标准数值进行对比分析, 进行变道效率、变道安全性、变道合理性等多方面的主成分分析考虑后得出最优解, 并且将此最优解转化程具体方案最终传输到控制执行模块最终实现变道的整个过程中对于车速和方向的把控。

值得一提的是, 整个变道的过程, 信息采集模块采用的负反馈系统, 而并不只是一次性的采集与决策, 与路况进行实时互动, 并且实时更新决策结果, 利用动态的分析使得车辆可以随时观测路面情况的变化以预防突发状况的发生, 使得无人驾驶汽车在实现自主变道时候的安全性和有效性得到极大的提升。

而实现上述功能的根本是在于决策系统和决策计算方法, 此决策系统使得之前所述的四个模块, 也就是用于最终进行真实操作车辆的主控制模块, 对信息进行实时采集的信息采集模块, 对于指令是否最终执行的控制执行模块和最终与信息进行事实交互并进行负反馈调整的信息反馈模块, 通过以上四个系统之间的相互交互, 最终计算出本车需要的转向角角度和加速度。

#### 4.4 自动驾驶中对红绿灯路口决策:

对于红绿等路口决策, 采用基于交通信息检测领域中目前最为权威的无人驾驶的动态失事路口甄别方法。此技术首先针对待处理图像中ROI进行切割, 并通过经验值对与所关心ROI (也即红绿灯高度相关区域) 之外的东西进行过滤。其次, 对标准的红绿灯模板进行建模作为对照组, 并绘制二维直方图作为HSV色彩空间的标准直方图方便进一步对比。对的待处理图像, 设置反转块投影进行搜索, 基于位置, 在YCBCR空间中对颜色指标进行进一步甄别。然后分别获得红、绿区域的坐标位置并进行比较, 根据红、绿的位置信息和智能车辆的车道信息决定是否行驶。本发明可以实时动态地检测红绿灯信息, 并应用于无人驾驶车辆。

要解决这个问题，首先必须能够从复杂的背景图片中获得红绿灯的大致位置，然后判断红绿灯的具体信息。目前，无人驾驶车辆的红绿灯识别方法主要包括神经网络等方法。该方法具有一定的检测精度，但受样品特性的影响。如果只有样本足以代表问题的特征，效果是明显的，而且样本的数量要适当，所以样本的选择就显得尤为重要。

对交通灯dst\_Cut2 的位置进行颜色识别。以上两步已经得到了红绿灯的准确位置，下一步就是识别颜色了。RGB空间也可用于颜色识别，但受光照等环境影响较大，因此转换到YCBCR空间的时间短，识别准确。提取符合交通灯颜色的cr通道值，对符合条件的区域进行二值化处理，得到绿灯二值化图bw\_green和红灯二值化图bw\_red。这样就可以对交通灯进行准确的颜色识别了。最后一步就是决策算法做出对应的正确的决策。

## 5 自动驾驶汽车决策算法的研究前景

### 5.1 尚可提升之处

当前的决策算法不能完全保证所有道路问题的安全，决策模块不仅需要考虑汽车的安全性和舒适性，以确保尽快到达目标位置，还需要确保乘客在复杂路况下的安全，甚至恶意损害乘客的利益和生命。因此，一方面，决策模块需要对驾驶计划进行长期规划；另一方面，它需要预测周围车辆和行人的行为。此外，无人驾驶中的决策模块对安全性和可靠性有严格的要求。现有的无人决策模块通常是根据规则构建的。虽然基于规则的构建可以处理大多数驾驶情况，但基于规则的决策系统不可能列举驾驶过程中可能发生的各种意外情况。我们需要一个自适应系统来处理驾驶环境中的各种紧急情况，因此这也是决策算法需要改进的地方。只有乘客的人身安全得到保障，无人驾驶才会被更多人接受。

道路的安全问题往往并不是我们在进行建模的时候就可以考虑完全的。无人驾驶在未来如果想要取得更加长足的进步，就要更加综合的只有将环境的不确定性纳入到决策控制算法中，才能更好的保证乘客的人身安全。

### 5.2 发展的有利之处

首先，无人驾驶汽车有政策倾斜。2022 年 1 月，《现代综合交通运输体系建设“十四五”规划》提出，要合理发展自动驾驶和车路协同等出行服务，鼓励在港口、物流园区等限制区开展自动驾驶测试应用，推进智能公交、智能停车、智能安检等发展。由于政策利好，无人驾驶行业有望迅速发展。

第二个方面是 5G 技术的加入。随着技术的发展,无人驾驶的决策也越来越精准,加入 5G 算法以后,无人驾驶也会越来越兴起。同时其也会和新能源技术产生配合,更加速无人驾驶技术的发展。

第三个方面是应用物流推动无人驾驶产业的发展。目前,对“物流无人”的需求不断增加。展望未来,物流领域的无人驾驶应用将进一步拓展,无人物流可以有效的避免病毒的传播,可以有效地保护居民的生活物品的供给,所以物流会不断推动无人驾驶产业的发展。无人驾驶汽车的产业化如今还存在着各种问题,但无人驾驶汽车本质上是建立在车辆主动安全技术和智能技术逐步升级的基础上。未来的交通系统可能会发生我们现在无法想象的变化,目前无人驾驶技术的发展重点仍然是辅助驾驶,从而提高人工驾驶的安全性。

#### 参考文献

[1] Zhang J P, Li L, Zhu Y. Research progress in behavioral decision making methods of unmanned vehicles based on reinforcement learning[J]. Electronics Technology, 2021, 34(05):66-71. (in Chinese)

张佳鹏, 李琳, 朱叶. 基于强化学习的无人驾驶车辆行为决策方法研究进展[J]. 电子科技, 2021, 34(05):66-71.

[2] Huang Z Q, Qv Z W, Zhang J, Zhang Y X, Tian R. End-to-end unmanned decision-making based on deep reinforcement learning[J]. Acta Electronica Sinica, 2020, 48(09):1711-1719. (in Chinese)

黄志清, 曲志伟, 张吉, 张严心, 田锐. 基于深度强化学习的端到端无人驾驶决策[J]. 电子学报, 2020, 48(09):1711-1719.

[3] Liu, Q., Li, X., Yuan, S. and Li, Z., 2021, September. Decision-making technology for autonomous vehicles: Learning-based methods, applications and future outlook. In 2021 IEEE International Intelligent Transportation Systems Conference (ITSC) (pp. 30-37). IEEE.

[4] W. Schwarting, J. Alonso-Mora and D. Rus, "Planning and decision-making for autonomous vehicles" in Annual Review of Control Robotics and Autonomous Systems, 2018.

- [5] Li, L., Ota, K. and Dong, M., 2018. Humanlike driving: Empirical decision-making system for autonomous vehicles. *IEEE Transactions on Vehicular Technology*, 67(8), pp.6814–6823.
- [6] Gallardo, N., Gamez, N., Rad, P. and Jamshidi, M., 2017, June. Autonomous decision making for a driver-less car. In *2017 12th System of Systems Engineering Conference (SoSE)* (pp. 1–6). IEEE.
- [7] Schwarting, W., Alonso-Mora, J. and Rus, D., 2018. Planning and decision-making for autonomous vehicles. *Annual Review of Control, Robotics, and Autonomous Systems*, 1, pp.187–210.
- [8] Li, G., Yang, Y., Zhang, T., Qu, X., Cao, D., Cheng, B. and Li, K., 2021. Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios. *Transportation research part C: emerging technologies*, 122, p.102820.
- [9] Nie, J., Zhang, J., Ding, W., Wan, X., Chen, X. and Ran, B., 2016. Decentralized cooperative lane-changing decision-making for connected autonomous vehicles. *IEEE access*, 4, pp.9413–9420.
- [10] Li, G., Yang, Y., Li, S., Qu, X., Lyu, N. and Li, S.E., 2022. Decision making of autonomous vehicles in lane change scenarios: Deep reinforcement learning approaches with risk awareness. *Transportation research part C: emerging technologies*, 134, p.103452.
- [11] Hubmann, C., Becker, M., Althoff, D., Lenz, D. and Stiller, C., 2017, June. Decision making for autonomous driving considering interaction and uncertain prediction of surrounding vehicles. In *2017 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1671–1678). IEEE.
- [12] E. Yurtsever, J. Lambert, A. Carballo and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies", *IEEE Access*, vol. 8, pp. 58443–58469, 2020.
- [13] W. Schwarting, J. Alonso-Mora and D. Rus, "Planning and decision-making for autonomous vehicles" in *Annual Review of Control Robotics and Autonomous Systems*, 2018.

- [14] D. Isele, "Interactive Decision Making for Autonomous Vehicles in Dense Traffic", 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pp. 3981–3986, 2019.
- [15] C. Vallon, Z. Ercan, A. Carvalho and F. Borrelli, "A machine learning approach for personalized autonomous lane change initiation and control", 2017 IEEE Intelligent Vehicles Symposium (IV), pp. 1590–1595, 2017.
- [16] R. Tami, B. Soualmi, A. Doufene, J. Ibanez and J. Dauwels, "Machine learning method to ensure robust decision-making of AV s", 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pp. 1217–1222, 2019.
- [17] F. Codevilla, M. Müller, A. López, V. Koltun and A. Dosovitskiy, "End-to-End Driving Via Conditional Imitation Learning", 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 4693–4700, 2018.
- [18] Mnih V, Kavukcuoglu K, Silver D, et al. Human — level control through deep reinforcement learning [J]. Nature, 2015, 518( 7540) : 529 — 533.
- [19] S. Aradi, "Survey of Deep Reinforcement Learning for Motion Planning of Autonomous Vehicles" IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 2, pp. 740–759, Feb. 2022