Intelligent driving system for all-domain intelligent logistics and autonomous driving

FORUM8

13th CPWC

Intelligent mobility in all domains

Overview

This project proposes a truck autonomous driving decision system based on a Vision-Language Model (VLM), with Mini-InternVL2-DA-BDD (DriveLM direction) as the core model, validated in the high-fidelity UC-win/Road simulation platform.

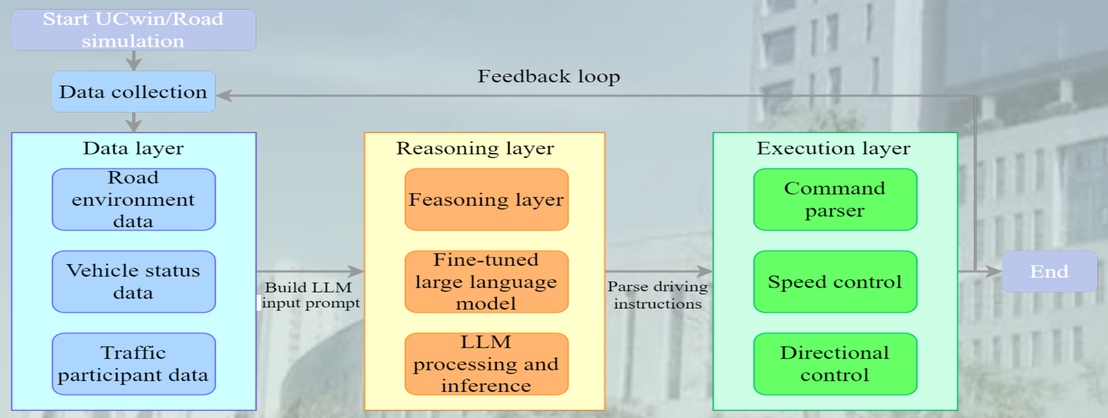
Objectives:

- Improve intelligent decision-making and safety in complex logistics scenarios.
- Achieve efficient inference under constrained compute (RTX 4060 8GB) while maintaining near large-model performance.
- Support unmanned upgrades across ports, industrial parks, and highway–urban distribution logistics.

System Architecture & Algorithm

Algorithm architecture design

This project designed a novel layered autonomous driving decision-making algorithm that integrates simulation environment data, the reasoning capabilities of a large language model (LLM), and vehicle control strategies to achieve safer and more intelligent autonomous driving. Three-layer autonomous driving decision system combining sim data, LLM reasoning, and control, refined via feedback loops.



Core Algorithm - Large Visual Language Model

Core Model：It uses the Mini-InternVL2-DA-BDD large-scale visual language model, which has efficient visual language understanding capabilities and is suitable for complex driving scenarios.

Data augmentation：Combined with the DriveGPT4 dataset, it simulates human driving decisions, generates high-quality driving behavior data, and improves strategy generation capabilities.

Real-time processing：Combined with the UC-win/Road simulation environment, it processes vehicle and environment information in real time, generates natural language driving suggestions and converts them into control instructions.

Balanced performance： Under limited hardware conditions, this model runs efficiently on the RTX 4060 (8GB) graphics card, ensuring the feasibility and effectiveness of the algorithm

Function Advantages

UC-win/Road-based project integrating the Mini-InternVL2-DA-BDD VLM to simulate real driving and evaluate performance in complex traffic; includes environment setup, data collection, model integration, and experiments for autonomous trucks.

Use APIs like GetVehicleInformation() to get vehicle status, turn it into natural-language prompts for Mini-InternVL2 to generate driving recommendations, then parse them into control signals to drive the truck—leveraging VLMs for smarter, more adaptive decisions.

Function	datatype	explain
Timestamp	datetime	The timestamp of the recorded data indicates the specific time when the data was collected.
ID	int	A vehicle's unique identifier, used to identify and distinguish between different vehicles.
Description	string	The description of the vehicle, including type, color, etc., such as "user's vehicle" or "blue car"
Pos_X	float	The vehicle's X coordinate position in 3D space (meters).
Pos_Y	float	The vehicle's Y coordinate position in 3D space (meters).
Pos_Z	float	The Z coordinate position of the vehicle in 3D space (meters), usually representing the height of the vehicle.
Yaw	float	The vehicle's yaw angle (radians), which represents the angle between the vehicle's heading and the X-axis.
Pitch	float	The vehicle's pitch angle (radian) affects the vehicle's fore-aft tilt.
Roll	float	The vehicle's roll angle (radian) affects the vehicle's left and right tilt.
Dir_X	float	The X-axis component of the vehicle's travel direction.
Dir_Y	float	The Y-axis component of the vehicle's travel direction.
Dir_Z	float	The Z-axis component of the vehicle's travel direction.
BodyPitch	float	The pitch angle of the vehicle body (in radians).
BodyRoll	float	The roll angle of the vehicle body (in radians).

UC-win/Road vehicle information output function table (partial)

Application Prospects

This technology improves efficiency in logistics and transportation, reduces labor costs, and reduces traffic accidents. At ports, it enables precise operations of driverless container trucks. It can also be applied in mines and construction sites, driving the intelligent and unmanned transformation of the freight industry.