Research on an Autonomous Vehicle Driven by a Large Language Model

Intern: SHIH YU CHENG

Project Manager: YEOH JIAN HERN

Contents

1. Introduction

This project aims to explore the feasibility of using a Large Language Model (LLM) for autonomous driving, specifically to pilot a vehicle within a simulated campus environment. Our primary objective is to enable an LLM to interpret real-time road conditions and generate precise driving commands, including *throttle*, *steering*, and *brake* controls. To achieve this, we have selected the *Phi-4-multimodal-instruct* model, which natively supports multi-modal inputs like images and text. This allows it to directly process visual information from the simulation to make driving decisions. The entire framework is built upon the CARLA simulator, which provides both a realistic campus environment and a vehicle dynamics system.

2. Methodology

The core methodology revolves around an interactive loop between the LLM and the CARLA simulator, supplemented by a fine-tuning mechanism analogous to Reinforcement Learning (RL).

2.1 System Framework

The system's operational flow is depicted in Figure 1. First, the vehicle in the CARLA simulator collects real-time road information using virtual cameras and LiDAR sensors. This multi-modal data is passed to the Phi-4 model. The LLM analyzes this information and generates corresponding numerical commands for throttle, steering, and braking. Finally, these commands are sent back to CARLA to control the vehicle's movement.

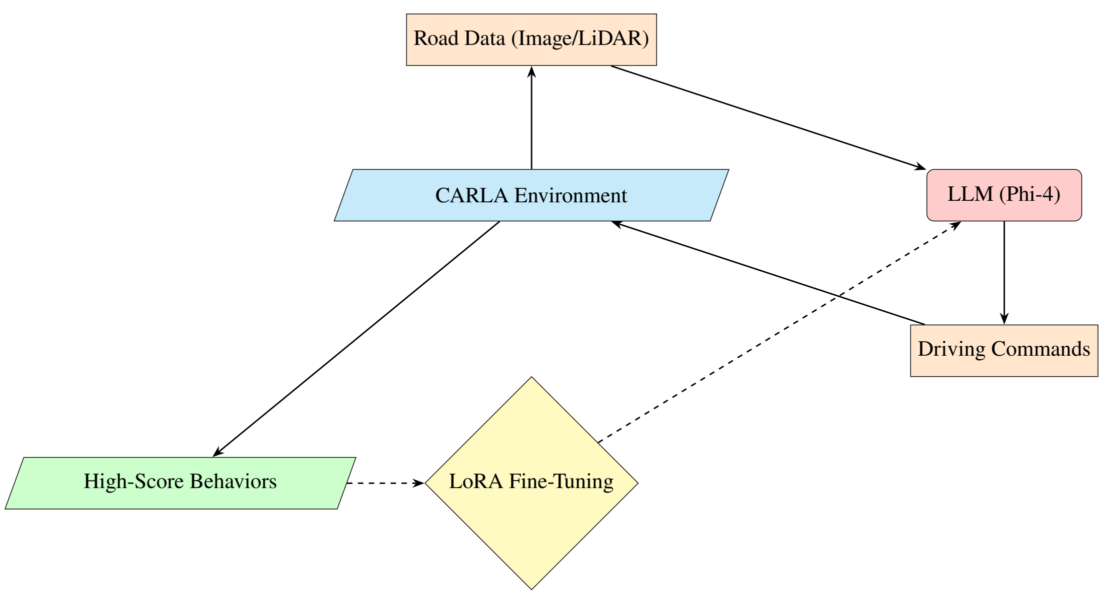


Figure 1: System Interaction and Model Fine-Tuning Flowchart

2.2 Model Fine-Tuning

We employ a strategy similar to reinforcement learning to optimize the LLM's driving capabilities. Specifically, we record a series of driving behaviors and score them using a performance matrix. High-scoring driving data is then selected and used to fine-tune the Phi-4 model's parameters using *Low-Rank Adaptation (LoRA)*. The goal is for the model to learn and replicate these proficient driving patterns.

2.3 Vehicle Longitudinal Dynamics Model

To enable the LLM to issue more accurate throttle commands, we referenced the work of Salvador Dominguez et al. to model the vehicle's longitudinal dynamics [1]. Their research proposes a method of using spline curves to fit propulsion, friction, and braking forces based on real-world response data. Figure 2 shows the propulsion force model from their paper, which describes the complex relationship between force and speed under different throttle inputs. We will use this model as a basis to teach the LLM to understand and apply the relationship between a desired acceleration and the required throttle value.

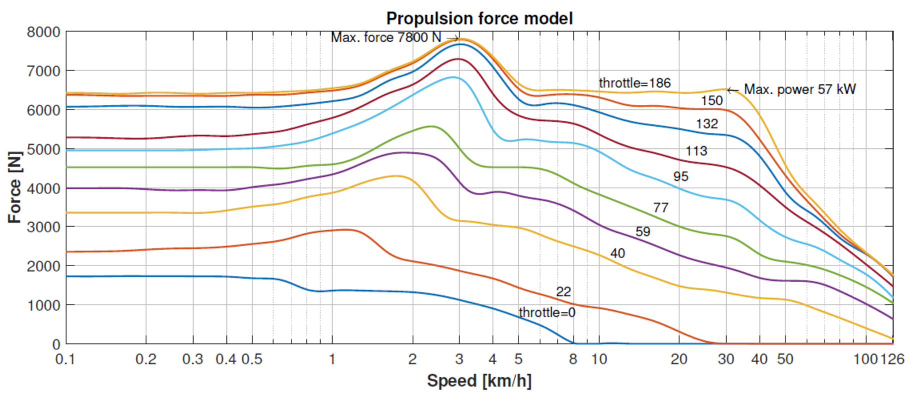
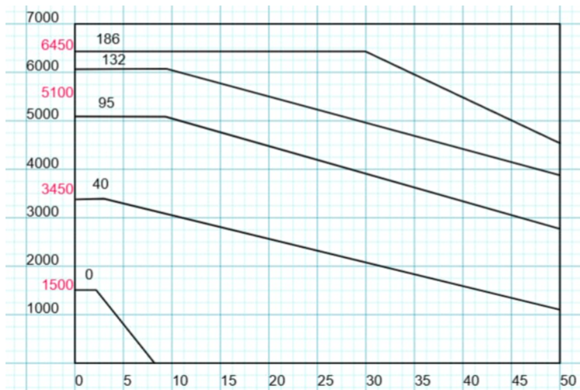


Figure 2: Propulsion force model of the electric vehicle from the reference paper [1].

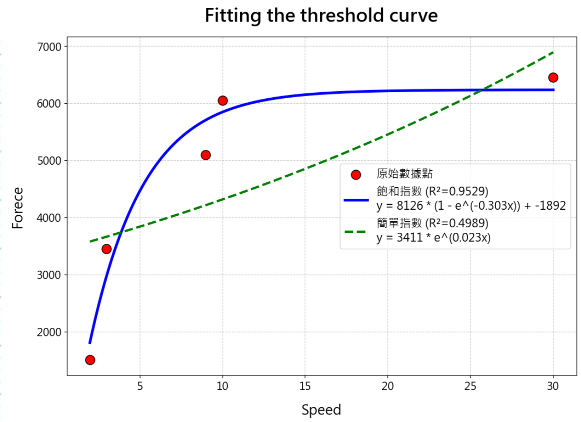
3. Preliminary Results

Currently, the project has successfully established the required campus driving scene in the CARLA simulator. The present focus is on validating the core capabilities of the Phi-4 model and on prompt engineering.

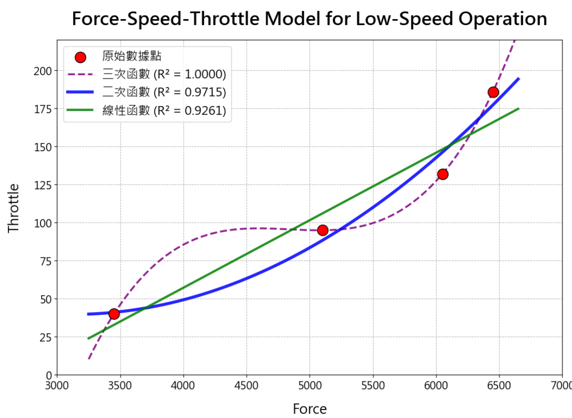
We are actively working to enable the LLM to master the aforementioned vehicle dynamics model, turning it into an effective tool for calculating the necessary throttle value to achieve a desired acceleration. To do this, we have simplified the complex model from the reference paper and performed a preliminary analysis by fitting functions to the data, as shown in the combined Figure 3.



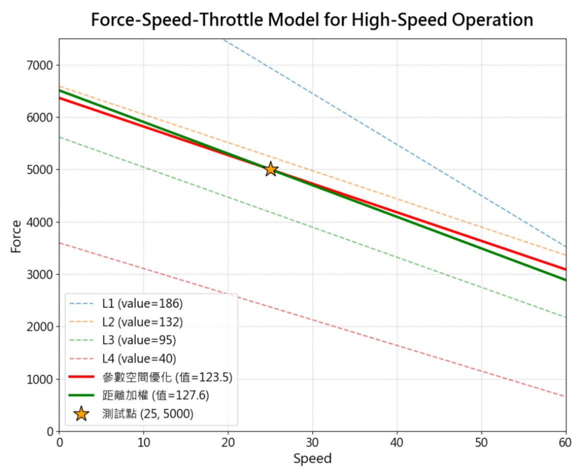
(a) Simplified Throttle-Force-Speed Model.



(b) Exponential Function Fitting Analysis.



(c) Fitting for lower throttle values.



(d) Fitting for higher throttle values.

Figure 3: Analysis of the simplified vehicle dynamics model. (a) shows the overall model, and (b) shows the threshold curve that determines which fitting model to use. Model (c) is applied to the lower-speed region (points above the curve), while model (d) is applied to the higher-speed region (points below it).

4. Future Works

Based on our current progress, future research will proceed in the following directions:

- **Complete Model Integration:** Continue with prompt engineering to ensure the LLM can reliably and accurately use the throttle-acceleration model.
- **Implement Fine-Tuning Loop:** Activate the LoRA-based fine-tuning process, using high-performance data collected in the simulation to continuously improve the model.
- **Comprehensive Evaluation:** Conduct extensive driving tests in the CARLA campus scene to assess the final model's autonomous driving capability, safety, and smoothness.

5. References

References

[1] Dominguez, S., Garcia, G., Hamon, A., & Frémont, V. (2020, October). Longitudinal Dynamics Model Identification of an Electric Car Based on Real Response Approximation. In *2020 IEEE Intelligent Vehicles Symposium (IV)* (pp. 398-404). IEEE.