AdaDrive: Self-Adaptive Slow-Fast System for Language-Grounded Autonomous Driving

Supplementary Material

1. Comparisons of Language-Grounded Driving Paradigms

We present the comparisons of our method with existing language-grounded driving paradigms across three key properties:

- Decoupling between LLM and planner operations (**D**)
- Asynchronous execution at different frequencies (**As**)
- Adaptive scheduling autonomy (**Ad**)

Table 1. Comparison results of our and existing paradigms for language-grounded driving, in terms of three critical properties.

Generation	Method	**D**	**As**	**Ad**
I	[3], [5]	✗	✗	✗
II	[1], [4]	✓	✓	✗
III	Ours	✓	✓	✓

The results are exhibited in Table 1. Specifically, LMDrive [3] and AD-H [5] employ a sequential paradigm where the LLM reasoning and trajectory prediction are coupled in a single forward pass, resulting in unavoidable inference delays. We classify this design paradigm as Generation I of language-grounded driving paradigms. Subsequently, the second generation of driving frameworks [1, 4] emerge with parallel and asynchronous architectures, enabling the integration of low-frequency LLM reasoning with high-frequency trajectory planning. However, their fixed asynchronous scheduling lacks flexibility, compromising the system's ability to handle complex or emergency situations. In contrast, our proposed AdaDrive framework not only implements an asynchronous slow-fast architecture but also incorporates adaptive mechanisms for LLM activation and integration, enabling dynamic response to varying driving conditions.

2. Extended Details of Benchmarks

Dataset: The training data consists of 64K samples from the official language-driven autonomous driving dataset [3], collected in the CARLA [2] environment across 8 towns. Each sample includes:

- Multi-sensor data: RGB images from four views (front, rear, left, right) and LiDAR scans.
- Navigation instructions: Natural language commands guiding the vehicle's movement, such as "*it's imperative to make a right turn at the next traffic signal*" and "*keep going straight until you reach the next junction*".

Table 2. Ablation of different slow-fast fusion strategies on the LangAuto-Short benchmark.

Strategy	Weight	DS ↑	RC↑	IS↑
Trajectory Averaging	1:1	63.7	75.8	0.80
Feature Fusion	1:1	66.4	74.2	**0.85**
Connector-H	adaptive	**70.6**	**85.3**	0.81

Table 3. Ablation of using different LLMs on the LangAuto-Short benchmark.

LLM	Params	DS ↑	RC↑	IS↑
TinyLLaMa	1.1B	70.6	85.3	0.81
LLaVA	7B	**75.9**	**89.3**	**0.84**

Benchmark: We evaluate our model using the standard LangAuto benchmarks [3] on the CARLA simulator. The benchmarks comprise three tracks with distinct route lengths:

- LangAuto: routes longer than 500 meters.
- LangAuto-Short: routes between 150 and 500 meters.
- LangAuto-Tiny: routes shorter than 150 meters.

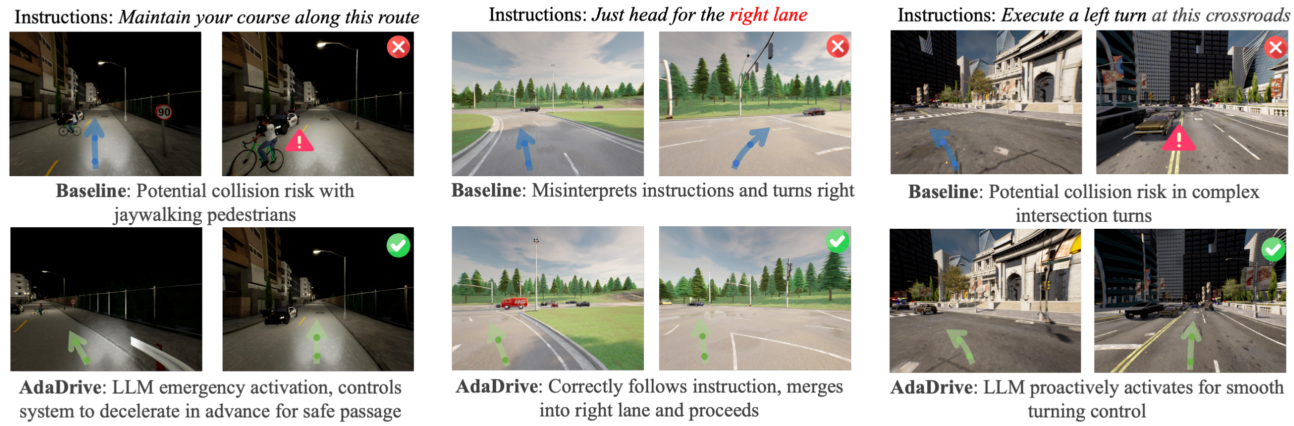
3. More Experiments

Analysis of Different Slow-Fast Fusion Strategies. We compare our Connector-H controlled dynamic LLM contribution scaling strategy against two full-weight baselines: 1:1 trajectory averaging and 1:1 feature fusion. Table 2 demonstrates that our adaptive fusion strategy achieves superior driving performance.

Analysis of Using Different LLMs: Beyond TinyLLaMa, we also implement our method with LLaVA-7B (using LoRA). As shown in Table 3, incorporating LLaVA further improves the DS score to 75.9%, demonstrating the scalability of our approach.

4. Visualizations

We provide some visualization results in Fig. 1, demonstrating representative scenarios where AdaDrive activates its LLM component adaptively in response to: (a) long-tailed emergency scenarios (e.g., jaywalking pedestrians), (b) challenging or ambiguous instruction comprehension and reasoning contexts (e.g. multi-stage navigation) and (c) complex road conditions. These results intuitively demonstrate how AdaDrive balances robustness and efficiency through dynamic LLM scheduling.



(a) **Emergency scenarios** (b) **Challenging or ambiguous instructions** (c) **Complex road conditions**

Figure 1. Qualitative comparison of the Baseline and AdaDrive, showing AdaDrive's dynamic LLM activation for enhanced driving safety.

References

[1] Yuan Chen, Zi-han Ding, Ziqin Wang, Yan Wang, Lijun Zhang, and Si Liu. Asynchronous large language model enhanced planner for autonomous driving. In *European Conference on Computer Vision*, pages 22–38. Springer, 2025. 1

[2] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017. 1

[3] Hao Shao, Yuxuan Hu, Letian Wang, Steven L Waslander, Yu Liu, and Hongsheng Li. Lmdrive: Closed-loop end-to-end driving with large language models. *arXiv preprint arXiv:2312.07488*, 2023. 1

[4] Xiaoyu Tian, Junru Gu, Bailin Li, Yicheng Liu, Chenxu Hu, Yang Wang, Kun Zhan, Peng Jia, Xianpeng Lang, and Hang Zhao. Drivevlm: The convergence of autonomous driving and large vision-language models. *arXiv preprint arXiv:2402.12289*, 2024. 1

[5] Zaibin Zhang, Shiyu Tang, Yuanhang Zhang, Talas Fu, Yifan Wang, Yang Liu, Dong Wang, Jing Shao, Lijun Wang, and Huchuan Lu. Ad-h: Autonomous driving with hierarchical agents. *arXiv preprint arXiv:2406.03474*, 2024. 1