LACNS: Language-Assisted Continuous Navigation in Structured Spaces

Rutong Peng, Yiqing Zhang, Yi Yang , Mengyin Fu

Beijing Institute of Technology

Abstract

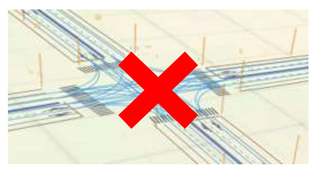
Motivations:

- Leveraging **Assisted Language** : integrate readily available assistive technologies
- Overcoming HD Map Limitations: prioritize real-time **perception-based** navigation.
- Embodied AI: employ **ChatGPT 4** and **LLAVA** for richer, context-aware scene understanding

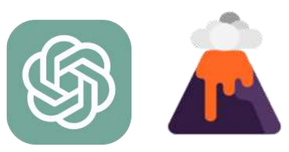
(a) Assisted Language

Contributions:

- Propose a language navigation framework, LACNS
- Synergize the capabilities of LLM and VLM
- Validate LACNS using the Carla simulator



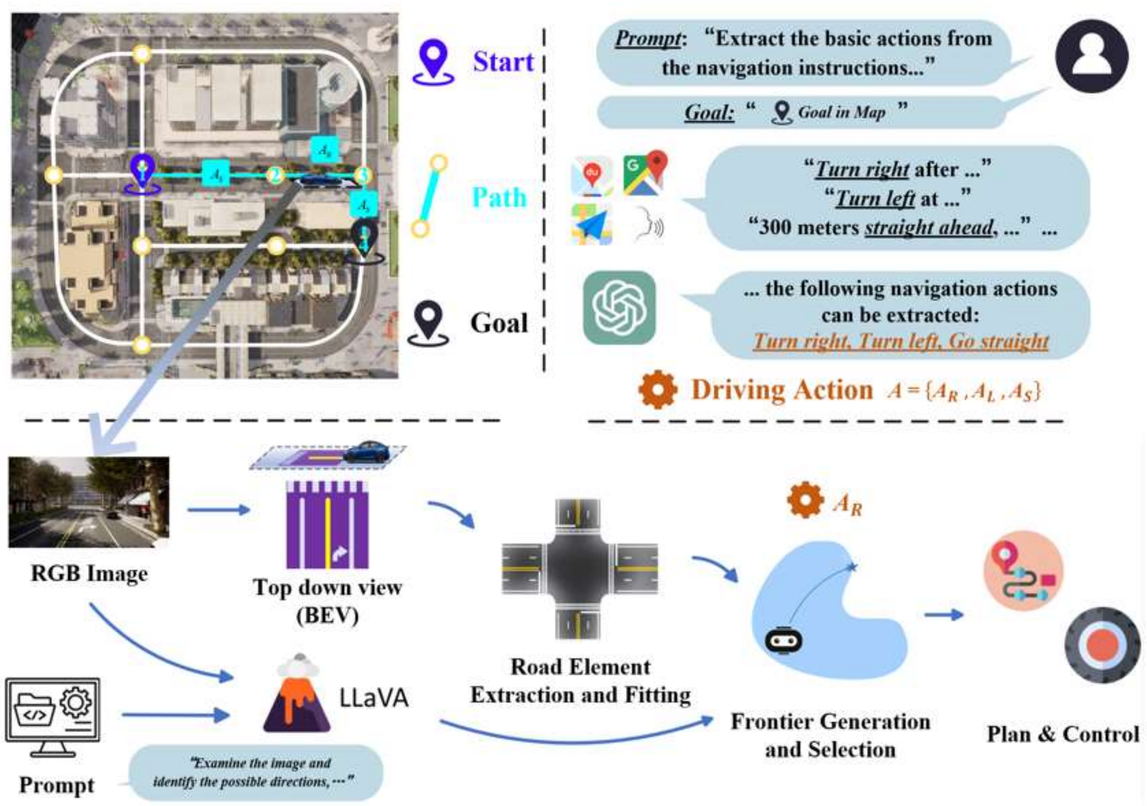
(b) No HD Map



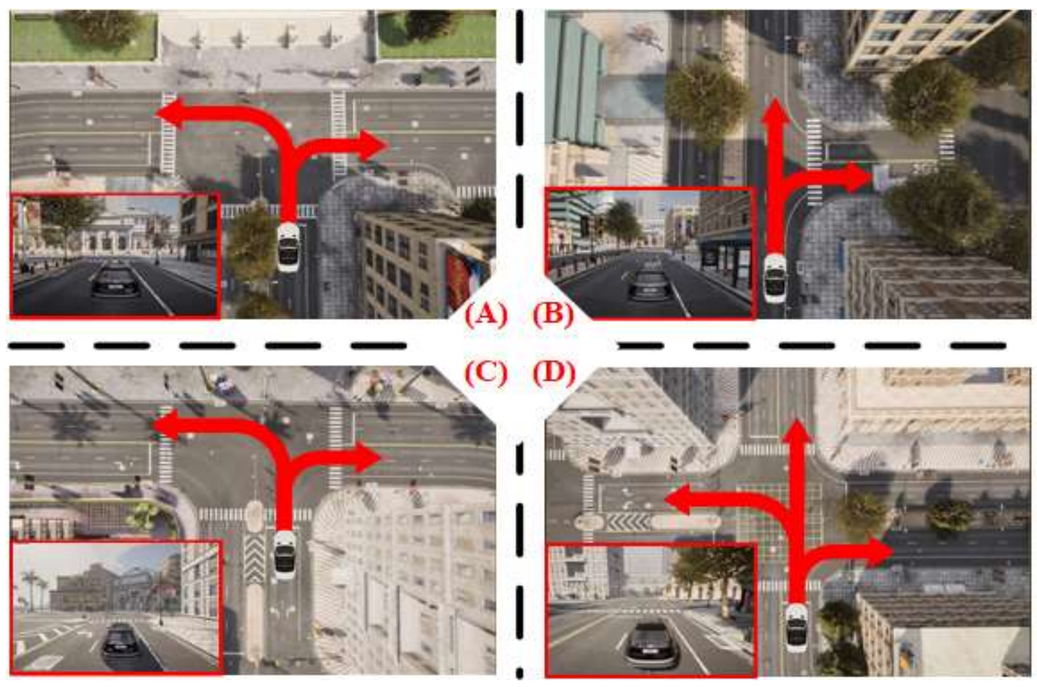
(c) Embodied Understanding

Methodology

- Process language instructions by **LLM**
- Generate a **BEV** map
- Detect intersections and assign a score by **VLM**
- Identify potential navigation boundaries
- Plan and control



Visualization Results



We present **dynamic** experiments deploying LACNS in the **Carla** environment, enabling vehicles to follow drivable instructions in a structured, urban, continuous space.

Interested? Scan the QR code, watch the video, and learn more.

Quantitative Experiment

Experimental results under different road conditions with individual instruction and sequential instruction.

SR: Success Rate CR: Compliance Rate NC: Number of completed instructions
CC: Number of instructions completed in compliance with traffic rules

Intersection	Go Straight		Turn Left		Turn Right	
	SR	CR	SR	CR	SR	CR
Road 1	×	×	0.92	0.69	0.85	0.69
Road 2	1.00	0.92	×	×	0.92	0.92
Road 3	×	×	0.77	0.62	0.85	0.62
Road 4	0.92	0.77	0.85	0.85	0.77	0.77

Instruction Sequence	Action Num	SR	CR	NC	CC
Sequence 1	3	0.62	0.54	2.38	2.23
Sequence 2	9	0.46	0.31	7.08	6.54
Sequence 3	9	0.38	0.15	7.08	5.38

Individual instruction experiments showed that the vehicle is stable when performing a single driving maneuver, and sequential instruction experiments showed that LACNS maintained this stability during complex navigation tasks.