LightPlanner: Unleashing the Reasoning Capabilities of Lightweight Large Language Models in Task Planning

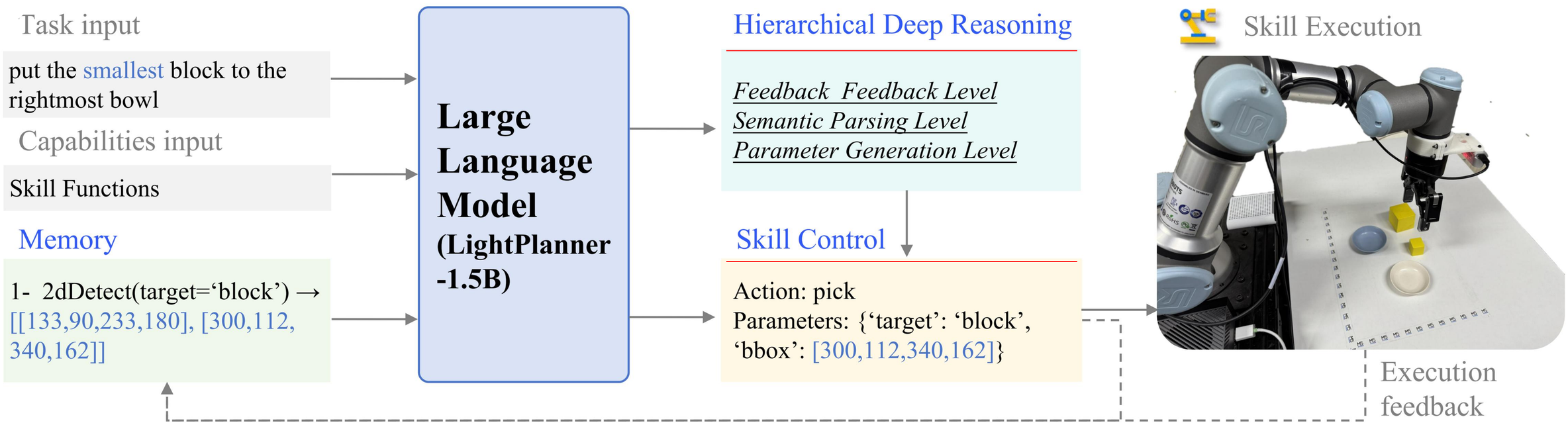
Weijie Zhou, Manli Tao, Chaoyang Zhao, Honghui Dong, Ming Tang, Jinqiao Wang

Beijing Jiaotong University
Institute of Automation, Chinese Academy of Sciences

Introduction & Motivation

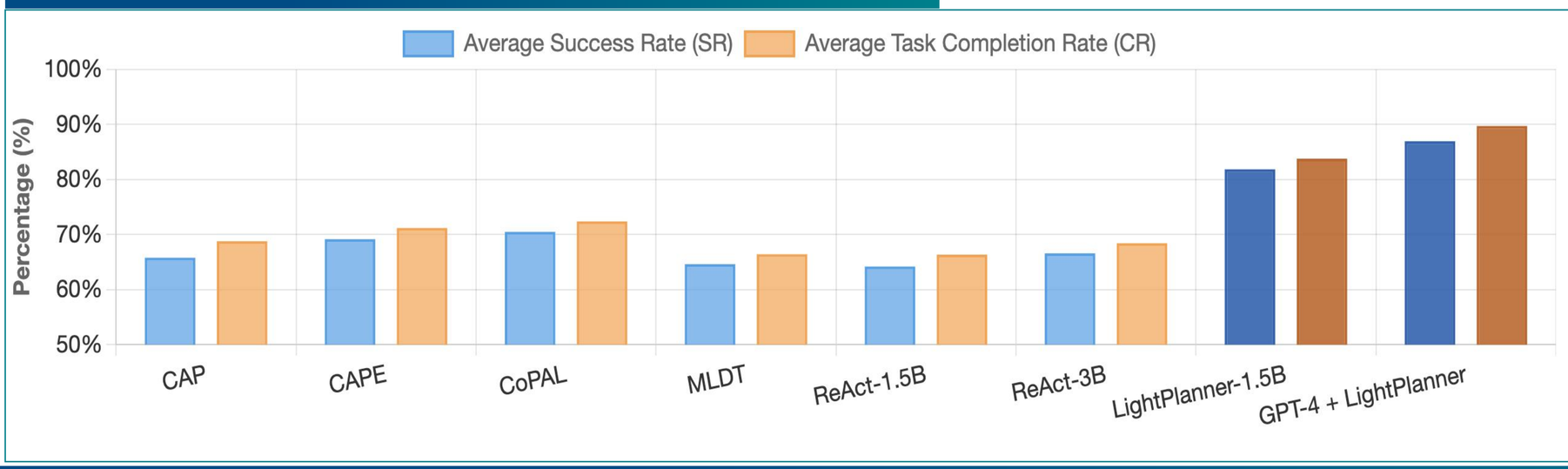
- **Goal:** Enable lightweight Large Language Models (LLMs) to perform complex robotic task planning on edge devices.
- **Challenge:** Standard lightweight LLMs (<3B parameters) excel at efficiency but fail in tasks requiring multi-step, dynamic logical reasoning (e.g., "pick the largest block").
- **Problem:** Existing methods rely on high-performance LLMs (like GPT-4) or use fixed skill templates that lack flexibility for dynamic commands.
- **Our Solution:** LightPlanner, a novel planner that enhances the reasoning of small LLMs to achieve high success rates in complex tasks

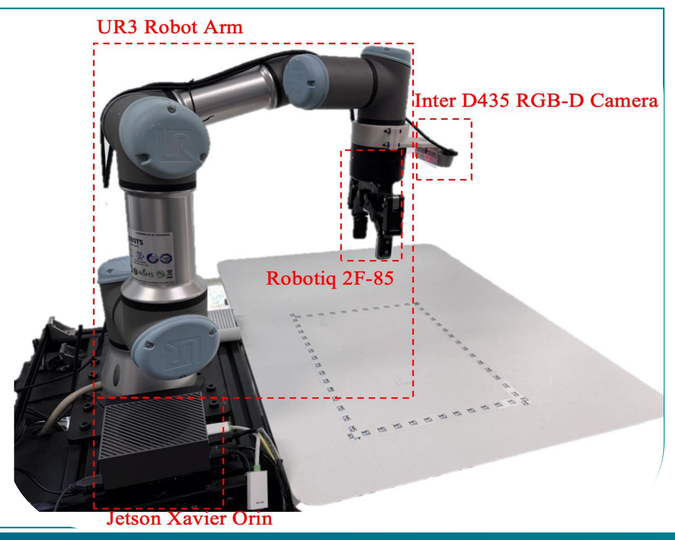
The LightPlanner Framework



Execution Feedback Level: Okay, I've located two blocks. One has a bounding box of [130, 90, 230, 180] and the other is [300, 110, 340, 160].
Semantic Parsing Level: Alright, my mission is to grab the smallest block and place it in the rightmost bowl. I've identified two blocks and now I'm going for the smaller one. This is exactly what I should be doing according to my task goal.
Parameter Generation Level: Okay, I need to pick the smallest block. I've already detected two blocks with bounding boxes [130, 90, 230, 180] and [300, 110, 340, 160]. To determine the smallest, I need to calculate the area of each bounding box.
The area of the first block is (233 - 133) * (180 - 90) = 100 * 90 = 9000.
The area of the second block is (340 - 300) * (162 - 112) = 40 * 50 = 2000.
Since 2000 is smaller than 9000, the second block with the bounding box [300, 112, 340, 162] is the smallest. Now I can proceed with the 'pick' action, using the bounding box of the smallest block.

Key Results





Weijie Zhou
PhD. Candidate
BJTU & CASIA
Elonzhou99@gmail.com
https://jetteezhou.github.io

My research bridges Embodied AI and MLLMs to build intelligent agents that perceive, reason, and act in the physical world.
First-author publications at IROS'25, CVPR'25, & NIPS'25.
Contributed to the TAIROS platform at Tencent Robotics X Lab, earning the Outstanding Intern Award (1 of 5 recipients).
Seeking collaborations in Embodied Cognition and Decision-making.