Zesen Lyu

⚲ Hangzhou China   ✉ lv_zesen@163.com   ☎ +86 151 7543 2001   💬 lzs_2001

Research Interests

Multimodal Large Model, Reasoning, Embodied AI, AI4S

Education

University of Chinese Academy of Sciences (UCAS)   Sep. 2023 – Jul. 2026
MS in Artificial Intelligence   GPA: 3.86/4.0 (Top 10%)

Hebei University of Technology (HEBUT)   Sep. 2019 – Jul. 2023
BS in Software Engineering   GPA: 3.57/4.0 (Top 15%)

Publications and Manuscripts

1. **Jigsaw-Puzzles: From Seeing to Understanding to Reasoning in Vision-Language Models** ⧉ (First Author)   Mar. 2025 – Aug. 2025
EMNLP 2025 Main
Keywords: VLM, Spatial Reasoning, Benchmark

2. **Structural fold mining- and deep learning-guided domain recombination (SMAL-DR) to explore the protein fitness landscape across evolution barriers** (Co-first Author)   Sep. 2024 - Jul. 2025
Nature Biotechnology
Under Review
Keywords: AI for Science, Protein Design.

Experience

Intent Understanding and Reasoning for Vertical Domains   Aug. 2025 – Dec. 2025
Research Intern   **Alibaba**
- Researching optimization strategies for embedding models.

Jigsaw-Puzzles: A Benchmark for Spatial Reasoning in VLMs   Mar. 2025 – Aug. 2025
Research Intern   **Zhejiang Lab**
- Motivated by the research paradigm of verifiable reward-based reinforcement learning for enhancing mathematical and programming capabilities, proposed a scalable and difficulty-controllable Jigsaw Puzzles benchmark in the multimodal domain.
- Designed five sub-tasks to systematically evaluate the spatial perception, understanding, single-step, and multi-step reasoning capabilities of large multimodal models, revealing a significant gap between current multimodal models and human performance in complex multi-step spatial reasoning.

Developing novel functional proteins through functional domain reprogramming   Sep. 2024 – Jul. 2025
Research Intern   **Zhejiang Lab**
- Collaborated with a postdoctoral researcher to design and implement a structural and functional domain mining and substitution pipeline for Cas9 proteins. Integrated structural segmentation algorithms and structural similarity search tools to construct a Cas9 domain database. Successfully identified multiple chimeric proteins with confirmed activity—some outperforming the wild-type—and trained an activity prediction model based on 70 wet-lab results to enable automated ranking and selection of candidate chimeras.

Open-World Planar Grasping Assisted by Large Multimodal Models   Jul. 2024 – Sep. 2024
Algorithm Intern   **WuBa Intelligent Tech.**
- During the internship, developed two demonstration pipelines for open-world planar grasping under the guidance of senior engineers. The first pipeline integrates YOLO-World for object detection, SAM for segmentation, and GraspNet for grasp pose estimation, enabling standard planar grasping.
- The second pipeline introduces multimodal foundation models and Set of Mark technology to replace traditional detection and segmentation modules, leveraging large models' planning capability to achieve sequential grasping and complex manipulations in open-world scenarios.

Awards and Honors

Academic Scholarship, University of Chinese Academy of Sciences   2024 – 2025 & 2023 - 2024
Outstanding Student Leader Award, University of Chinese Academy of Sciences   2024 – 2025
Merit Student Award, University of Chinese Academy of Sciences   2024 – 2025 & 2023 - 2024

Language and Skills

Languages: Chinese(Native), English(CET-6)
Skills: LLM Training, Finetune, Python, PyTorch, Git, Shell