Ychen Zhu

13862305183 | zhuyuchenno1.github.io |
zhych2023@shanghaitech.edu.cn | ZhuYuChenNO1
Master Student - ShanghaiTech University

Research Interest

My current research interests include 1. Spatial Intelligence, 2. Multimodal Large Language Models (LLM/MLLM), 3. Open-world Visual Understanding. Currently, I am focusing on equipping multimodal large language models with spatial reasoning abilities, particularly in 3D scene understanding, embodied navigation, and object grounding.

Publications

- **Rethinking Query-based Transformer for Continual Image Segmentation**
  Yuchen Zhu*, Cheng Shi*, Dingyou Wang, Jiajin Tang, Zhengxuan Wei, Yu Wu, Guanbin Li, Sibei Yang†*
  - **Accepted by IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2025**
  - We revisit query-based transformers from a continual learning perspective and propose a novel framework for continual image segmentation, enhancing model plasticity and mitigating catastrophic forgetting.
  - Code and models released for reproducibility: SimCIS
- **Plain-Det: A Plain Multi-Dataset Object Detector**
  Cheng Shi, **Yuchen Zhu***, Sibei Yang†*
  - **Accepted by European Conference on Computer Vision (ECCV) 2024**
  - We propose Plain-Det, a simplified end-to-end object detector designed to scale with both dataset size and diversity. Plain-Det leverages CLIP priors to resolve label conflicts and employs dynamic sampling strategies to handle joint training over datasets with varying sizes. Furthermore, it performs de-biasing on text embeddings to enhance classification accuracy. As the dataset scales up, Plain-Det effectively scales up in capability.
  - Code and models released for reproducibility: Plain-Det
- **Sim-DETR: Unlock DETR for Temporal Sentence Grounding**
  Jiajin Tang, Zhengxuan Wei*, **Yuchen Zhu**, Cheng Shi, Guanbin Li, Liang Lin, Sibei Yang†*
  - **Accepted by International Conference on Computer Vision (ICCV) 2025**
  - We proposes Sim-DETR, which resolves DETR's query conflicts in temporal sentence grounding via self-attention adjustments and query-frame alignment, unlocking superior performance and faster convergence.
- **Unified Mitigation of LVLM Hallucinations across Alignment Formats**
  Jiaye Qian, Ge Zheng*, **Yuchen Zhu**, Sibei Yang†*
  - **In Submission**

Education

- **ShanghaiTech University** *Sep. 2023 - Now*
  Master of Computer Science Shanghai, China
  - GPA: 3.33/4.00 (Major GPA: 3.45/4.00)
- **Nanjing Institute of Technology** *Sep. 2018 - June 2022*
  B.Eng. in Automation Nanjing, China
  - Major Courses: Automatic Control, Microprocessors, Digital Electronic Technology, Mechanics ...
  - Grade: 3.1/4.0

Projects

- **Undergraduate Thesis: Design of a Two-Wheeled Self-Balancing Vehicle Based on STM32** *2022*
  C++, Stm32, PID Control, Keil, MPU6050, PWM, Embedded Systems
  - Hardware: A two-wheeled self-balancing vehicle built around STM32F103C8T6, MPU6050, TB6612 and ultrasonic sensors, with complete circuit and mechanical design.
  - Software & Control: Cascade PID enables dynamic balance, velocity tracking, straight-line driving and autonomous obstacle avoidance.
- **Undergraduate Course Project: keyword recommendation multi-task IO server** *2021*
  C++, Linux, TCP/IP
  - Designed and implemented a multi-task I/O keyword recommendation server using the Reactor framework and thread pool. The server processes client queries by retrieving and ranking similar words from a combined Chinese-English dictionary using edit distance, cppjieba segmentation, and LRU-based caching for optimized performance.

Honors and Awards

- **ShanghaiTech Outstanding Student** *Dec. 2024*
- **CET-6 certificate** *Dec. 2019*
- **Scholarship awarded by the School of Automation** *Dec. 2018, 2019*