BOSHEN XU

website: xuboshen.github.io

(+86) 13881861005

boshenx@ruc.edu.cn

EDUCATION

Renmin University of China   *Sep. 2023 - now*
Ph.D in Computer Science   Advisor: Qin Jin
University of Electronic Science and Technology of China   *Sep. 2019 - Jun. 2023*
B.S in School of Computer Science and Engineering   Advisor: Qin Jin
GPA: 3.97/4.00(90.41/100), ranked 2/65

RESEARCH INTERESTS

egocentric computer vision and embodied AI.

PROJECTS

Human-Object Relation Understanding   *Nov. 2022- Jul. 2024*

- **Egocentric Vision-Language Pretraining (arxiv 2024)**. We discover that current Egocentric Vision-Language Models (EgoVLM) are lack of open-vocabulary EgoHOI recognition ability on our constructed benchmark EgoHOIBench. To address this issue, we propose EgoNCE++, an asymmetric contrastive learning pretraining objective to pretrain the EgoVLMs.

- **Egocentric Hand-Object Interaction (ACM MM 2023)**. The egocentric data is expensive to fetch. We propose a visual prompt-based method to effectively pre-train model with third-person videos to achieve viewpoint adaptation and viewpoint generalization to Ego-HOI recognition.

- **Human-Object Interaction Detection (CVPR 2023)** The HOI task has long been plagued by the lack of data and supervision due to the large number of possible combinations of verbs and objects in real life. Models trained on limited data can only predict HOIs within a set of fixed categories. We therefore propose an open-category HOI pre-training model with specially designed proxy tasks that can well generalize to novel interaction categories.

Object Relation Understanding in 3D Space.   *Aug. 2023- Nov. 2023*

- **Autonomous 3D Object Assembly (arxiv 2024)** We addresses the combinatorial explosion challenge, where the number of possible combinations rises beyond exponentially with increasing parts, by leveraging weak constraints from assembly sequences, effectively reducing the solution space's complexity.

PUBLICATIONS

1. Liangyu Chen, Zihao Yue, **Boshen Xu**, Qin Jin. Unveiling Visual Biases in Audio-Visual Localization Benchmarks. In ECCV AVGenL Workshop, 2024

2. **Boshen Xu**, Ziheng Wang*, Yang Du*, Zhinan Song, Sipeng Zheng*, Qin Jin. EgoNCE++: Do Egocentric Video-Language Models Really Understand Hand-Object Interactions? In arxiv, 2024

3. **Boshen Xu**, Sipeng Zheng, and Qin Jin. SPAFormer: Sequential Part Assembly with Transformers. In arxiv, 2024

4. **Boshen Xu**, Sipeng Zheng, and Qin Jin. POV: Prompt-Oriented View-agnostic Learning for Egocentric Hand-Object Interaction in the Multi-view World. In ACM MM, 2023

5. Sipeng Zheng, **Boshen Xu**, and Qin Jin. Open-category human-object interaction pre-training via language modeling framework. In CVPR, 2023

SERVICE

- Conference Reviewer: ICLR, ACM MM, ACCV.

AWARDS AND SCHOLARSHIPS

Provincial Outstanding Graduate.	*2023*
National Second Prize in China Undergraduate Mathematical Contest in Modeling.	*2021*
Tencent Special Scholarship.	*2021*
National Scholarship for Undergraduates.	*2020*

Oct 2, 2024