SIJIN CHEN

Fudan University | H: +86 187 1792 9716 | csjch3cook@gmail.com
Homepage | GitHub | Google Scholar

About Me

I am currently a final-year Master's student in Artificial Intelligene at Fudan University (Sep. 2021 - Jun. 2024, expected). I am supervised by Prof. Tao Chen, and am fortunate to work closely with Dr. Hongyuan Zhu. Before this, I obtained my Bachelor's degree in Data Science and Big Data Technology also from Fudan University (Sep. 2017 - Jun. 2021).

My long-term research goal is to develop vision-language systems that possess the capacity to **comprehend**, **reason**, and **envision** the physical world. Outside my research, I love sports and music.

Research Interests

Multi-modal Learning, Vision and Language, Large Language Models, 3D Scene Understanding.

Education

Masters in Artificial Intelligence (GPA 3.56/4.00, rank 30/351) Sep. 2021 - Jun. 2024
Fudan University. Supervised by Prof. Tao Chen. Shanghai, China
Bachelor in Data Science and Big Data Technology (GPA 3.20/4.00) Sep. 2017 - Jun. 2021
Fudan University. Shanghai, China

Publications and Pre-prints (google scholar)

- LL3DA: Visual Interactive Instruction Tuning for Omni-3D Understanding, Reasoning, and Planning.
Sijin Chen, Xin Chen, Chi Zhang, Mingsheng Li, Gang Yu, Hao Fei, Hongyuan Zhu, Jiayuan Fan, Tao Chen.
[Pre-print | project | paper | github]
[Summary]: An LLM 3D Assistant responds to visual and text interactions in complex 3D scenes.
- Vote2Cap-DETR++: Decoupling Localization and Describing for End-to-End 3D Dense Captioning.
Sijin Chen, Hongyuan Zhu, Mingsheng Li, Xin Chen, Peng Guo, Yinjie Lei, Gang Yu, Taihao Li, Tao Chen.
[Pre-print | paper | github]
[Summary]: Decoupled feature extraction and task decoding for 3D Dense Captioning.
- End-to-End 3D Dense Captioning with Vote2Cap-DETR.
Sijin Chen, Hongyuan Zhu, Xin Chen, Yinjie Lei, Gang Yu, Tao Chen.
[CVPR 2023 | paper | github | youtube]
[Summary]: We address 3D Dense Captioning as a set prediction problem with parallel decoding.
- M3DBench: Let's Instruct Large Models with Multi-modal 3D Prompts.
Mingsheng Li, Xin Chen, Chi Zhang, **Sijin Chen**, Hongyuan Zhu, Fukun Yin, Gang Yu, Tao Chen.
[Under Review | paper]
[Summary]: A large scale dataset querying 3D LLMs with text, 2D, and 3D prompts.
- WI3D: Weakly Incremental 3D Detection via Visual Prompts.
Mingsheng Li, **Sijin Chen**, Shengji Tang, Hongyuan Zhu, Xin Chen, Fukun Yin, Tao Chen.
[Under Review | paper]
[Summary]: Introduce novel classes to a pre-trained 3D object detector with 2D foundation models.

Projects

- **Language for 3D Scenes.** Aug. 2021 - Jun. 2024
Proposed a set-to-set solution for 3D dense captioning, accepted by CVPR 2023, won the Scan2Cap challenge at ICCV 2023, and submitted an advanced Vote2Cap-DETR++ to T-PAMI, under review. Presented **LL3DA**, a large language 3D assistant responding to diverse human interactions in complex 3D environments, under review. Put forward **M3DBench**, a large-scale 3D vision language dataset covering 327k lines of annotations for 10 tasks related to 3D perception, understanding, reasoning, and planning, under review.
- **Class-Incremental 3D Detection.** Apr. 2023 - Dec. 2023
Proposed WI3D, enabling a trained 3D detector to detect new categories with images, under review.

- **Earlier Projects.** Before Sep. 2021
  Self-Supervised Pre-training on 3D Point Clouds. Developed a self-supervised learning algorithm that learns global- and patch-level contrastive representations for 3D point clouds.
  A Smart Advertisement Display System. Developed a human perception system that detects faces, recognizes facial expressions, estimates eye gaze, age, and gender for advertisement recommendation.

Scholarships and Awards

First place winner of the Scan2Cap Challenge at ICCV 2023. Oct. 2023
National Scholarship (rank 1/46). Sep. 2023
Second prize of the Scholarship for Outstanding Students at Fudan University. Sep. 2022
Award for the Scholarship for Outstanding Students at Fudan University. Sep. 2021
Second prize of the Scholarship for Outstanding Students at Fudan University. Jun. 2021

Invited Talks

Winner presentation of **the Scan2Cap Challenge** at ICCV 2023. [slides | youtube] Oct. 2023
Paper presentation at **the workshop for advances in 3D vision**, VALSE 2023. Apr. 2023

Skills

Languages:	Chinese (native), English (proficient), Shanghai dialect
Programming:	Python, R, C, Matlab, SQL
Tools:	PyTorch, Blender, Visual Studio, Spyder, Jupyter Notebook

References

Prof. Tao Chen Supervisor Fudan University eetchen@fudan.edu.cn