

Tianhe Ren

✉ rentianhe@idea.edu.cn

🐙 Github

🌐 Google Scholar

🌐 Homepage

I'm primarily interested in researching **vision foundation models, object detection and segmentation, and multi-modal learning**. I'm also passionate about **open-source projects** in AI community. The research work and open-source projects I'm involved in have garnered over **25.0K** stars on Github.

Employment History

- 2022 – Now 📌 **Computer Vision Engineer**, International Digital Economy Academy (IDEA), Computer Vision and Robotics Center.
Supervisor: Prof. Lei Zhang.
- 2021 – 2022 📌 **Computer Vision Engineer**, OneFlow, Vision Group.
Supervisor: Prof. Jinhui Yuan.

Education

- 2017 – 2021 📌 **Xiamen University**, China.
Bachelor, GPA: 3.59/4.00, Average Score: 86.48/100.00
Major: School of Information Science and Engineering
Supervisor: MAC Lab, Prof. Rongrong Ji.

Research and Project Highlight

- 2023 - Now 📌 **Grounded-Segment-Anything**: Marrying Grounding-DINO with Segment Anything & Stable Diffusion & Recognize Anything - Automatically Detect, Segment and Generate Anything.
Character: **Project Lead & First Author**
Github Stars: **15.3 K**.
Grounded-SAM combines the strong open-set detector Grounding-DINO with promptable segmentation model (SAM) for detecting and segmenting arbitrary regions with users' textual inputs. **Grounded-SAM** is widely applied in various influenced work such as Florence-2, Emu-Edit, Task-Matrix and so on.
- 2022 – Now 📌 **detrex**: detrex is a research platform for DETR-based object detection, segmentation, pose estimation and other visual recognition tasks.
Character: **Project Lead & First Author**
Github Stars: **2.0 K**.
detrex is a deep-learning library built upon detectron2 and mainly focus on the transformer-based detection algorithms. **detrex** supports over 15 mainstream detection transformer algorithms and further boosts their performance from 0.2 AP to 1.1 AP by optimizing both model and training hyper-parameters.
- 2023 - Now 📌 **Grounding-DINO**: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection
Character: **Project Lead & Main Contributor**
Github Stars: **6.9 K**.
Grounding-DINO combines the strong DINO detector with large-scale grounded pre-training which can detect any regions based on the user's textual inputs.

Selected Publications

The full paper list is available at my Google Scholar. (* Equal contributions. List order is random.) And my full project list is available at my Github.

Total Google Scholar Citations: **2000+**, Total Github Stars: **25000+**.

1. **DINO-X: A Unified Vision Model for Open-World Object Detection and Understanding**
Tianhe Ren*, Yihao Chen*, Qing Jiang*, Zhaoyang Zeng*, Yuda Xiong*, Wenlong Liu, Zhengyu Ma, Junyi Shen, Yuan Gao, Xiaoke Jiang, Xingyu Chen, Zhuheng Song, Yuhong Zhang, Hongjie Huang, Han Gao, Shilong Liu, Hao Zhang, Feng Li, Kent Yu, Lei Zhang
Tech report, Nov. 2024
2. **Grounding DINO 1.5: Advance the "Edge" of Open-Set Detection**
Tianhe Ren*, Qing Jiang*, Shilong Liu*, Zhaoyang Zeng*, Wenlong Liu, Han Gao, Hongjie Huang, Zhengyu Ma, Xiaoke Jiang, Yihao Chen, Yuda Xiong, Hao Zhang, Feng Li, Peijun Tang, Kent Yu, Lei Zhang
Tech report, May. 2024
3. **detrex: Benchmarking Detection Transformers**
Tianhe Ren*, Shilong Liu*, Feng Li*, Hao Zhang*, Ailing Zeng, Jie Yang, Xingyu Liao, Ding Jia, Hongyang Li, He Cao, Jianan Wang, Zhaoyang Zeng, Xianbiao Qi, Yuhui Yuan, Jianwei Yang, Lei Zhang
Tech report, May. 2023
4. **Grounded SAM: Assembling Open-World Models for Diverse Visual Tasks**
Tianhe Ren*, Shilong Liu*, Ailing Zeng, Jing Lin, Kunchang Li, He Cao, Jiayu Chen, Xinyu Huang, Yukang Chen, Feng Yan, Zhaoyang Zeng, Hao Zhang, Feng Li, Jie Yang, Hongyang Li, Qing Jiang, Lei Zhang
International Conference on Computer Vision (ICCV) Demo Track, 2023
Citation: 170, Github star: 15.3K
5. **Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection**
Shilong Liu, Zhaoyang Zeng, **Tianhe Ren**, Feng Li, Hao Zhang, Jie Yang, Chunyuan Li, Jianwei Yang, Hang Su, Jun Zhu, Lei Zhang
European Conference on Computer Vision (ECCV), 2024
ECCV 2024 The 1st Most Influential Paper
Citation: 1300+, Github star: 6.9K
6. **Detection Transformer with Stable Matching**
Shilong Liu*, **Tianhe Ren***, Jiayu Chen*, Zhaoyang Zeng, Hao Zhang, Feng Li, Hongyang Li, Jun Huang, Hang Su, Jun Zhu, Lei Zhang
International Conference on Computer Vision (ICCV), 2023
7. **Cheap and Quick: Efficient Vision-Language Instruction Tuning for Large Language Models**
Gen Luo, Yiyi Zhou, **Tianhe Ren**, Shengxin Chen, Xiaoshuai Sun, Rongrong Ji
Conference on Neural Information Processing Systems (NeurIPS), 2023
8. **You Only Segment Once: Towards Real-Time Panoptic Segmentation**
Jie Hu, Linyan Huang, **Tianhe Ren**, Shengchuan Zhang, Rongrong Ji, Liujuan Cao
Computer Vision and Pattern Recognition (CVPR), 2023
9. **T-Rex: Counting by Visual Prompting**
Qing Jiang, Feng Li, **Tianhe Ren**, Shilong Liu, Zhaoyang Zeng, Kent Yu, Lei Zhang
Tech report, Nov. 2023
10. **TRAR: Routing the Attention Spans in Transformers for Visual Question Answering**
Yiyi Zhou, **Tianhe Ren**, Chaoyang Zhu, Xiaoshuai Sun, Jianzhuang Liu, Xinghao Ding, Mingliang Xu, Rongrong Ji
International Conference on Computer Vision (ICCV), 2021

Selected Projects

- 2023 - Now **Grounded-Segment-Anything:** Marrying Grounding-DINO with Segment Anything & Stable Diffusion & Recognize Anything - Automatically Detect, Segment and Generate Anything.
Github Stars: 15.3 K.
- Grounding-DINO:** Marrying DINO with Grounded Pre-Training for Open-Set Object Detection
Github Stars: 6.9 K.
- 2022 - 2023 **detrex:** detrex is a research platform for DETR-based object detection, segmentation, pose estimation and other visual recognition tasks.
Github Stars: 2.0 K.
- 2021 - 2022 **LiBai:** A Toolbox for Large-Scale Distributed Parallel Training based on the OneFlow Deep Learning Framework.
Github Stars: 371.

Professional Services

- Conference Reviewer **Computer Vision in the Wild (CVinW)**, on CVPR 2024 Workshop
European Conference on Computer Vision (**ECCV**), 2024
Conference on Neural Information Processing Systems (**NeurIPS**), 2024
International Conference on Learning Representations (**ICLR**), 2025
Artificial Intelligence and Statistics (**AISTATS**), 2025
Computer Vision and Pattern Recognition (**CVPR**), 2025
- Organizer **Computer Vision in the Wild (CVinW)**, on CVPR 2023 Workshop

Skills

- Programming **Python, \LaTeX**
- Programming Tools **PyTorch, PyTorch-Lightning, Scikit-Learn, Git, Linux**