Xin LAI — 赖昕

xinlai@cse.cuhk.edu.hk +86 188 4604 4208 Age: 25 years old

Introduction

I am currently a final-year Ph.D student at CUHK, supervised by Prof. Jiaya Jia (IEEE Fellow). My research interests lie in 2D/3D Scene Understanding, Multi-modal Large Language Model, Vision Transformer, Data-efficient Learning. My representative works include LISA (enabling multi-modal LLMs with reasoning segmentation capabilities) and Stratified Transformer (a fully transformer-based 3D fundamental network). [Google Scholar]

Education Background

The Chinese University of Hong Kong (CUHK) — **Hong Kong, China**
Ph.D Student of Computer Science and Engineering — *Aug. 2020-Aug. 2024 (Expected)*
Supervisor: Prof. Jiaya Jia (IEEE Fellow).

Harbin Institute of Technology (HIT) — **Harbin, China**
Bachelor of Computer Science and Technology — *Sept. 2016-Jun. 2020*
GPA: 93.8/100, Ranking: 1st / 39 in Computer Engineering track

The University of Hong Kong (HKU) — **Hong Kong, China**
Exchange Student of Computer Science — *Sept. 2018-Jan. 2019*

Publications

1. **LISA: Reasoning Segmentation via Large Language Model**
   Xin Lai, Zhuotao Tian, Yukang Chen, Yanwei Li, Yuhui Yuan, Shu Liu, Jiaya Jia
   International Conference on Learning Representations (ICLR), 2024. (In submission) [Paper][Code]
   Over 800 GitHub Stars within two weeks! Invited to give a talk at **Adobe (California, US)** and other institutions. Reported by **top-tier technology media**. [量子位] [Twitter] [机器之心] [CVer]

2. **Mask-Attention-Free Transformer for 3D Instance Segmentation**
   Xin Lai, Yuhui Yuan, Ruihang Chu, Yukang Chen, Han Hu, Jiaya Jia
   International Conference on Computer Vision (ICCV), 2023. [Paper][Code]

3. **Spherical Transformer for LiDAR-based 3D Recognition**
   Xin Lai, Yukang Chen, Fanbin Lu, Jianhui Liu, Jiaya Jia
   Computer Vision and Pattern Recognition (CVPR), 2023. [Paper][Code]

4. **Stratified Transformer for 3D Point Cloud Segmentation**
   Xin Lai, Jianhui Liu, Li Jiang, Liwei Wang, Hengshuang Zhao, Shu Liu, Xiaojuan Qi, Jiaya Jia
   Computer Vision and Pattern Recognition (CVPR), 2022. [Paper][Code]
   A pioneering fully transformer-based 3D network. Invited to give a talk at **Microsoft Research Asia (MSRA)**.

5. **DecoupleNet: Decoupled Network for Domain Adaptive Semantic Segmentation**
   Xin Lai, Zhuotao Tian, Xiaogang Xu, Yingcong Chen, Shu Liu, Hengshuang Zhao, Liwei Wang, Jiaya Jia
   European Conference on Computer Vision (ECCV), 2022. [Paper][Code]

6. **Semi-supervised Semantic Segmentation with Directional Context-aware Consistency**
   Xin Lai, Zhuotao Tian, Li Jiang, Shu Liu, Hengshuang Zhao, Liwei Wang, Jiaya Jia
   Computer Vision and Pattern Recognition (CVPR), 2021. [Paper][Code]

7. **Learning Context-aware Classifier for Semantic Segmentation**
   Zhuotao Tian, Jiequan Cui, Li Jiang, Xiaojuan Qi, **Xin Lai**, Yixin Chen, Shu Liu, Jiaya Jia

AAAI Conference on Artificial Intelligence (AAAI), 2023. [Paper]

8. **Generalized Few-shot Semantic Segmentation**
   Zhuotao Tian, **Xin Lai**, Li Jiang, Michelle Shu, Hengshuang Zhao, Jiaya Jia
   Computer Vision and Pattern Recognition (CVPR), 2022. [Paper][Code]

9. **Adaptive Perspective Distillation for Semantic Segmentation**
   Zhuotao Tian, Pengguang Chen, **Xin Lai**, Li Jiang, Shu Liu, Hengshuang Zhao, Bei Yu, Ming-Chang Yang, Jiaya Jia
   IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2022.

10. **Guided Point Contrastive Learning for Semi-supervised Point Cloud Semantic Segmentation**
    Li Jiang, Shaoshuai Shi, Zhuotao Tian, **Xin Lai**, Shu Liu, Chi-Wing Fu, Jiaya Jia
    International Conference on Computer Vision (ICCV), 2021. [Paper]

Experiences

- **Tencent ARC Lab**
  - Mentor: Dr. Yanpei Cao
  - Working on the intersection between Multi-modal Large Language Model and vision tasks.
- **SmartMore**
  - Mentor: Dr. Shu Liu
  - Working on 2D/3D scene understanding, semi-supervised, transfer, few-shot semantic segmentation.
  - Early exploration of fully transformer-based 3D fundamental network (Stratified Transformer).
  - Plug-and-play radial transformer module to enhance distant objects recognition (Spherical Transformer).
  - 1st on nuScenes & SemanticKITTI 3D semantic segmentation benchmark (Nov., 2022)

Open-source Projects

- **Sparse Transformer (SpTr)** [Code]
  A fast, memory-efficient, and easy-to-use library for sparse transformer with varying token numbers (e.g., window transformer for 3D point cloud). It is highly optimized by underlying CUDA code.
  Author: **Xin Lai** (the core contributor, contributions include optimization ideas, all code writing, testing, and time evaluation), Fanbin Lu, Yukang Chen

Awards

- Outstanding Reviewer — *by CVPR, 2023*
- Full Postgraduate Studentship — *2020-2024*
- Outstanding Graduate — *by HIT, 2020*
- National Scholarship (Top 1%) — *by Chinese Ministry of Education, 2018*
- Fung Scholarship — *by Victor and William Fung Foundation Ltd., 2018*
- Provincial Merit Student (Top 2%) — *2018*
- People's Scholarship — *by HIT, 2016-2018*

Invited Talks

- "LISA: Reasoning Segmentation via Large Language Model" at **Adobe (California, US)**
- "Stratified Transformer" at **Microsoft Research Asia (MSRA)** [Slides]
- "Spherical Transformer" at **TechBeat 2023** [Slides]
- "DecoupleNet" at **AI Times 2022** [Slides]

Activities & Teaching

- Reviewer for CVPR, ICCV, ECCV, NeurIPS, TIP, Pattern Recognition.
- Teaching Assistant: ENGG 1110: Problem Solving By Programming / CSCI 3251: Engineering Practicum