XUANYAO CHEN

RM501, BD3, No.2500 Songhuajiang RD ⋄ Hongkou District, Shanghai
(+86)198-2123-7398 ⋄ xuanyaochen19@fudan.edu.cn

EDUCATION

Fudan University *September 2019 - Present*
B.Eng. in Department of Electronic Engineering
GPA Rank(first 3 years): 1/113
Grade Highlights: Mathematical Analysis(A), Linear Algebra(A), Programming(A), College Physics(A), English Argumentative Writing(A), Engineering Mathematics(A), Probability and Mathematical Statistics(A), Fundamentals of Digital Logic(A), Analog Electronics Foundation Course(A), Computational Complexity(A), Signal and Communication Systems(A), Digital Signal Processing(A), Natural Language Processing(A).

RESEARCH EXPERIENCE

Massachusetts Institute of Technology Apr 2022 - Present
Research Intern in HAN Lab *Advisor: Prof. Song Han*

IIIS, Tsinghua University Mar 2021 - Oct 2022
Research Intern in MARS Lab *Advisor: Prof. Hang Zhao*

HONORS & AWARDS

- SenseTime Scholarship(30 undergrads all over China) Dec 2022
- Samsung Scholarship Oct 2021
- The First Price, Chinese Undergraduate Mathematics Competition(rank 6 in Shanghai) Apr 2021
- National Scholarship(rank 1st in class) Oct 2020
- Undergraduate Innovation Award, Fudan Oct 2020
- The Second Price, Chinese Undergraduate Mathematical Contest in Modeling Sep 2020

PUBLICATION

SparseViT: Revisiting Activation Sparsity for Efficient High-Resolution Vision Transformer
Xuanyao Chen, Zhijian Liu*, Haotian Tang, Li Yi, Hang Zhao, Song Han*
CVPR 2023

ViP3D: End-to-end Visual Trajectory Prediction via 3D Agent Queries
Junru Gu, Chenxu Hu*, Tianyuan Zhang, Xuanyao Chen, Yilun Wang, Yue Wang, Hang Zhao*
CVPR 2023

FUTR3D: A Unified Sensor Fusion Framework for 3D Detection
Xuanyao Chen, Tianyuan Zhang*, Yue Wang, Yilun Wang, Hang Zhao*

MUTR3D: A Multi-camera Tracking Framework via 3D-to-2D Queries
Tianyuan Zhang, Xuanyao Chen, Yue Wang, Yilun Wang, Hang Zhao
CVPR 2022. Workshop on Autonomous Driving.

What Makes Multimodal Learning Better than Single (Provably)
Yu huang, Chenzhuang Du*, Zihui Xue, Xuanyao Chen, Hang Zhao, Longbo Huang*
NeurIPS 2021.

RESEARCH PROJECTS

Efficient Vision Transformer Design
Supervisor: Prof. Song Han *April 2022 - Present*

- Explore an efficient image feature extraction backbone based on Swin Transformer to reduce MACCs and latency of inference, aiming to make model more practical and easier to deploy. It can achieve significant measured latency reduction compared to out efficient design.
- Achieve 1.5× latency speedup without accuracy loss. Our model can outperform other popular pruning methods.
- Accepted by CVPR 2023.

Multi-modal object detection in autonomous driving
Supervisor: Prof. Hang Zhao *July 2021 - Sep 2022*

- Propose a simple yet effective query-based method which utilizes 3D point clouds and RGB images to solve the object detection problem in autonomous driving.
- Demonstrate our method's generality and effectiveness in different combination of sensors, like 1/4/16/32 beams LiDAR, Radar with cameras.
- The sensor fusion framework enables low-cost autonomous driving, while only using a 4-beam LiDAR with cameras, our method achieves on par performance with state-of-the-art 3D detection model using a 32-beam LiDAR.

Multi-cameras multi objects tracking
Supervisor: Prof. Hang Zhao *Sep 2021 - Apr 2022*

- Propose a unified query-based framework for multi objects tracking in multi cameras, directly predicting object trajectories in 3D space and then projected to 2D images to sample image features.
- Out method achieves state-of-the-art results in nuScenes benchmark.

Multi-modal learning theory analysis
Supervisor: Prof. Hang Zhao *Apr 2021 - July 2021*

- Propose theoretical justifications in multi-modal influential empirical results to capture important qualitative phenomena observed in real multi-modal applications.
- Show that multi-modal learning does possess an appealing formal guarantee.
- Conduct several multi-modal experiments to support our theoretical analysis.

Supervised depth estimation
Supervisor: Prof. Hang Zhao *Apr 2021 - Mar 2021*

- Use PackNet model to predict per-pixel depth of a single image under the supervision of ground-truth depth from LiDAR.
- Introduce smoothness loss of predicted depth map for balancing overfitting in sparse LiDAR points and estimation accuracy.

SKILLS

Programming Languages C, C++, Python, Matlab, LaTex, JavaScript, HTML