

# Wenhao Yao

2205 Songhu Rd, Yangpu Dist, Shanghai

☎ (+86) 135 3520 0540 | ✉ [whyao23@m.fudan.edu.cn](mailto:whyao23@m.fudan.edu.cn) | 🌐 [william-yao-2000.github.io](https://github.com/william-yao-2000)

## EDUCATION

### Fudan University

Sep 2023 – Present

B.Sc., Computer Science and Technology, GPA 3.41

### South China University of Technology

Sep 2019 – Jun 2023

B.Eng., Computer Science and Technology, GPA 3.95

## RESEARCH EXPERIENCE

### DriveSuprim: Towards Precise Trajectory Selection for E2E Planning

Dec 2024 – June 2025

- Selection-based methods in end-to-end planning face several challenges: the inability to distinguish hard negative trajectories, the dominance of straight-driving trajectories, and binary trajectory labels that hinder optimization.
- Propose a coarse-to-fine refinement method to improve trajectory discrimination, and a pipeline combining rotation-based data augmentation with self-distillation to address directional bias and hard decision boundaries.
- Achieves state-of-the-art performance on both the NAVSIM and Bench2Drive datasets, with acceptable model size and inference efficiency.

### Multi-Modal Prototypes for Vast-Vocabulary Object Detection

May 2024 – Aug 2024

- The main challenge of vast vocabulary object detection is in enormous visual concept classification.
- Extract features from example images and detailed descriptions to build classifiers and distinguish visual concepts, and ensemble the logits of different classifiers to enhance performance.
- Surpass V3Det dataset SOTA results in both Vast Vocabulary and Open Vocabulary Object Detection settings.
- Reach perfect result in *V3Det Challenge 2024* (1st place in the Open Vocabulary Object Detection Track, and 2nd place in the Vast Vocabulary Object Detection Track).

### Unleashing General Mask Transformers for Scene Text Spotting

Sep 2023 – Aug 2024

- Build a simple scene text spotting pipeline based on the Mask Transformers, while preserving the pre-trained visual knowledge of the detection module.
- Extract text foreground features from segmented results through random sampling, and utilize a non-autoregressive Transformer to recognize text.
- Strengthen the synergy effect between text detection and recognition by the object query-text alignment.
- Achieve competitive and even better performance on ICDAR-2015, Total-Text and CTW1500 compared to previous state-of-the-art methods.

## TEACHING EXPERIENCE

### A Introduction to Artificial Intelligence, Fudan University

Mar 2024 – Jun 2024

*Teaching Assistant*

- Design and instruct lab projects on Convolutional Neural Networks.
- Prepare course notes on Unsupervised Learning, Representation Learning, and Games.

## AWARDS

**1st place**, CVPR 2025 NAVSIM v2 End-to-End Driving Challenge

May 2025

**1st place** in the Open Vocabulary Object Detection Track, and **2nd place** in the Vast Vocabulary Object Detection Track, CVPR 2024 V3Det Challenge

June 2024

**Second Prize**, National College Student Computer System Ability Competition – Compiler System Design

Aug 2021

## SKILLS

---

**Deep Learning Frameworks and Platforms:** Detectron2, Pytorch, Pytorch-Lightening

**Coding Skills:** C++, Python

**English:** CET-6 611

## OUTREACH

---

Monitor, 2019 Computer Science and Technology Class 1, South China University of Technology      Sep 2019 – Jun 2023

President, Guangdong Experimental High School Astronomy Association      Sep 2017 – Jun 2018