Haoyuan Yu

yuhaoyuan03@gmail.com | yu-haoyuan.github.io

Haoyuan Yu

Third-Year Undergraduate Student, **Hunan University**
Address: Changsha, Hunan, China
Last updated: March 2, 2025

Research Interests

- **Egocentric Vision Tasks:** Focus on hand and body tracking and segmentation within egocentric video, with potential applications in immersive VR gaming experiences.
- **Multimodal Audio-Visual Understanding:** Exploring the integration of audio and visual data to develop enhanced AI tools for video-based applications.
- **Egocentric Audio-Visual Fusion:** Investigating the fusion of egocentric audio and visual streams to enable embodied intelligence.

Education

Hunan University, BS in Computer Science   Sept 2022 – Jun 2026
                                                          (Expected)

- **Coursework:** Deep Learing(98), Computer Vision(94), Pattern Recognition(95), Data Mining (93)

Publications

Egocentric Speaker Diarization with Vision-Guided Clustering and Adaptive Speech Re-detection(ICASSP 2025)   Dec 2024

He Huang, *Haoyuan Yu*, Daibo Liu, Daibo Liu, Haowen Chen, Minjie Cai

Projects

Data Mining class project   project page

- This project involved the training and evaluation of a Hyper-YOLO model on the PTL-AI Furnas Dataset. It was completed as the final assignment for the 2022 Data Mining course at Hunan University.
- Tools Used: Python

Multimodal Image Retrieval task   tianchi.aliyun.com

- This project focused on an e-commerce image retrieval task, where the model retrieves relevant product images from a given pool based on natural language queries. We fine-tuned Chinese CLIP for this task and achieved a result of 7/71.

Egocentric Object Tracking

- We employed the STARK architecture, a baseline model for visual tracking, to address the Visual Queries 2D Localization challenge within the Ego4D dataset.

Technologies

Languages: Python, C++