Shengyu Hao

Email: shengyuhao@zju.edu.cn
Github: github.com/shengyuhao

EDUCATION

- **Zhejiang University** — Hangzhou, China
  Ph.D. in Software Engineering — *Sept. 2021 - Present*
  Advisor: *Prof. Gaoang Wang*

- **Beijing University of Posts and Telecommunications** — Beijing, China
  M.Sc. in Control Science and Engineering — *Sept. 2017 - Jun. 2020*
  Advisor: *Prof. Yanzhu Hu*

- **Shenyang Ligong University** — Shenyang, China
  B.Sc. in Automation — *Sept. 2013 - Jul. 2017*

PUBLICATIONS

- **Hao, S.**, Liu, P., Zhan, Y., Jin, K., Liu, Z., Song, M., Hwang, J., Wang, G. DIVOTrack: A novel dataset and baseline method for cross-view multi-object tracking in diverse open scenes. *International Journal of Computer Vision*, 2024.

- **Hao, S.**, Zhao, Z., Chai, W., Wang, G., et al. Ego3DT: Tracking All 3D Objects in Ego-Centric Video of Daily Activities. *ACM Multimedia*, 2024.

- Zhao, Z., Chai, W., Wang, X., Li, B., **Hao, S.**, Hwang, J., Wang, G., et al. See and think: Embodied agent in virtual environment. *European Conference on Computer Vision*, 2024.

- Cao, S., Chai, W., **Hao, S.**, Wang, G., et al. Difffashion: Reference-based fashion design with structure-aware transfer by diffusion models. *IEEE Transactions on Multimedia*, 2023.

- **Hao, S.**, Wang, G., Gu, R. Weakly supervised instance segmentation using multi-prior fusion. *Computer Vision and Image Understanding*, 2021.

RESEARCH EXPERIENCE

- **Understanding Dynamic 3D World in Ego-Centric Videos** — *Jan. 2024 - Present*
  Based on ego-centric videos, open-source perception models such as GLEE, SAM, and dust3r are used to construct 3D environments. Large language models assist agents in understanding daily activities, addressing challenges in dynamic scene tracking and 3D scene comprehension.

- **Cross-view Multi-Object Tracking** — *Sept. 2021 - Present*
  Cross-view multi-object tracking aims to associate objects between frames and camera views with geometric constraints and representation learning. It addresses cross-view association based on detection and multi-object tracking.

- **Embodied Agent in Virtual Environment** — *Jul. 2023 - Apr. 2024*
  This work presents an advanced embodied agent for Minecraft that integrates vision perception, language instruction, and code action to interpret visual information, decompose tasks, and execute skill actions.

- **Fashion Design by Diffusion Models** — *Dec. 2022 - Sept. 2023*
  This work presents DiffFashion, a novel diffusion model for reference-based fashion design that combines reference appearance and clothing images to generate new fashion images.

- **Weakly Supervised Instance Segmentation** — *Apr. 2021 - Oct. 2021*
  This work introduces a weakly supervised instance segmentation method that uses bounding box tightness and contour priors to improve mask prediction and achieve superior performance without mask-level labels.

TEACHING EXPERIENCE

- **ECE365: Data Science and Engineering**, Teaching Assistant — *Spring 2023*
- **CS412: Introduction to Data Mining**, Teaching Assistant — *Fall 2022*

AWARDS AND HONORS

- **Outstanding Graduate Student**, Zhejiang University — *2022 & 2024*
- **First-Class Academic Scholarship**, Beijing University of Posts and Telecommunications — *2017 - 2020*
- **Outstanding Graduate, Second-class Scholarship**, Shenyang Ligong University — *Dec. 2016*

ACADEMIC SERVICE

Reviewer, ICLR, Neurips, ACM MM, ECCV, CVPR, ICASSP, TCSVT, KBS