

KAI YE

Tel:(+86) 18559228224 | Email: kaiye1@link.cuhk.edu.cn | Website: <https://elvin-yk.github.io/kaiye.github.io/>

EDUCATION

The Chinese University of Hong Kong, Shenzhen

Jan. 2024 - Now

*Master of Philosophy, **Major:** Computer and Information Engineering, **Supervisor:** Prof. Rui Huang*

Lanzhou University

Sept. 2018 - July 2022

*Bachelor of Science, **Major:** Computer Science & Technology, **Supervisor:** Prof. Minqiang Yang*

PUBLICATION/PREPRINTS

- [1] **Kai Ye***, Yuhang Wu*, Shuyuan Hu, Junliang Li, Meng Liu, Yongquan Chen, Rui Huang. Gen2Real: Towards Demo-Free Dexterous Manipulation by Harnessing Generated Video. In *arXiv:2509.14178 Submitted to ICRA 2026*.
- [2] Junliang Li*, **Kai Ye***, Haolan Kang, Mingxuan Liang, Yuhang Wu, Zhenhua Liu, Huiping Zhuang, Rui Huang, Yongquan Chen. Grasp What You Want: Embodied Dexterous Grasping System Driven by Your Voice. In *arXiv:2412.10694 Major revision at Journal of Field Robotics*.
- [3] **Kai Ye**, Yinru Ye, Minqiang Yang, Bin Hu. Independent Encoder for Deep Hierarchical Unsupervised Image-to-Image Translation. In *arXiv:2107.02494*
- [4] Minqiang Yang, Yinru Ye, **Kai Ye**, Wei Zhou, Xiping Hu, Bin Hu. Retinal Vessel Segmentation in Medical Diagnosis using Multi-scale Attention Generative Adversarial Networks. *The Journal of Mobile Networks and Applications*
- [5] Minqiang Yang, Yinru Ye, **Kai Ye**, Xiping Hu, Bin Hu. Retinal Vessel Segmentation Using Multi-scale Generative Adversarial Network with Class Activation Mapping. *The 10th EAI International Conference on Wireless Mobile Communication and Healthcare, 2021*.
- [6] A kind of Eyeglasses Try-On System (**first student author**)
The National Invention Patent (publication number: CN112418138B)
- [7] Software Copyright: Lottery system for the right to use underground parking spaces (**first student author**)
- [8] Kexin Sun, Yuelan Xin, Yunliang Qi, Meng Lou, **Kai Ye**, Yinru Ye. CAGU-Net: Category Attention Guidance U-Net for Retinal Blood Vessel Segmentation. *17th International Conference on Computational Intelligence and Security (CIS). IEEE*.
- [9] Xinlong Chen, **Kai Ye**, WenCe Zhou. Application of artificial intelligence in new diagnostic and therapeutic pattern of pancreatic diseases and its advances. *Chinese Journal of Medical Physics, 2022*.

RESEARCH EXPERIENCE

Research Assistant, UAIS Lab, Lanzhou University

Mar. 2019 - Sept. 2021

Professors: Bin Hu, Minqiang Yang

Independent Encoder for Deep Hierarchical Unsupervised Image-to-Image Translation

- By restructuring the architecture of the conventional image translation network, the semantic inconsistency problem is effectively mitigated, while the overall performance of image translation is significantly enhanced. The proposed model was superior to the state-of-the-art model described in a CVPR paper at the time. [2]

Retinal Vessel Segmentation

- By introducing a multi-scale attention mechanism into adversarial neural networks and employing image stitching as a data augmentation technique, this study enhances the performance of retinal vessel segmentation. [3, 4]

Research of Eyeglasses Removal and Wearing Technology and Development of Virtual Try-On Eyeglasses Platform

- This study tackles the difficulty faced by highly myopic users who cannot clearly view the effect of new glasses in-store. It employs image-to-image translation to remove the user's current glasses and augmented reality to overlay the new glasses on the face, allowing virtual try-on without removing existing eyewear.

- Output: 1) National Invention Patent [5]; 2) National Key Innovation and Entrepreneurship Project; 3) The second prize at the provincial level and the special prize at the school level in the National College Student E-Commerce “Innovation, Creativity and Entrepreneurship” Challenge.

Research Assistant, ShenZhen Institute of Artificial Intelligence and Robotics

Jun. 2024 - now
Professors: *Huang Rui*

Multimodal Perception towards Embodied AI

- Participated in CVPR 2024 Autonomous Driving Competition on behalf of the laboratory and achieved **5th** place out of 64 teams in the **Multi-View 3D Visual Grounding** track. <https://opendrivelab.com/challenge2024>

Dexterous Hand Grasping for Embodied AI

- Developed a multi-modal grasping system that integrates voice commands and vision for precise target identification.
- Designed robust, human-inspired grasp strategies leveraging multi-finger coordination to handle complex objects.
- Output: 1) An academic paper [1]; 2) Excellence Award, 2024 Shenzhen Intelligent Robot Dexterous Hand Competition; 3) I was interviewed by CCTV during the 2024 China International Hi-Tech Fair.

WORK EXPERIENCE

Algorithm Engineer, YITU, Shanghai

Jul 2022 - Jun 2023

Autonomous Driving Perception

Jul 2022 - Aug. 2022

- Developed a replacement model for single-frame blockage detection using a radar-based scheme, integrated the data, tuned the models, and put the integrated model into a road test.

Autonomous Driving Motion Prediction

Sept. 2022 - Dec. 2022

- To improve the overall performance of the prediction model, I followed most of the relevant papers and implemented approaches, such as data augmentation, network architecture optimization, and loss function design, that were known to benefit the current network module. The final performance index was 10% lower than the benchmark experiment (lower is better) and the model successfully solved 44.05% of the road test problems in the actual scene.

Mapless Driving

Jan. 2023 - Jun 2023

- Mainly responsible for the data engineering part of road network structure prediction. Focus on designing experiments to improve performance from the perspective of data dimensions. While ensuring data accuracy and pipeline efficiency, explore the impact of data diversity, density, and other aspects on performance.

DEVELOPMENT EXPERIENCE

Eye tracker host computer development | *Main developer*

Dec. 2020 - Sept. 2021

- Developed a host computer system capable of reading and storing eye movement and facial expression data that are synchronously captured by a depth camera and eye tracker. This system has been used in the UAIS Lab.

Lottery system for the right of underground parking | *Independent developer*

Mar. 2019 - Oct. 2019

- Implemented a lottery system to allocate parking spaces for new apartments to teachers. The system has conducted an open lottery on site.

Multimodal open dataset MODMA website | *Main developer*

Nov. 2019 - Feb. 2020

- Developed a website to host a public dataset of emotions for UAIS labs. URL: <http://modma.lzu.edu.cn/>

Logistics robot | *Main developer*

Sept. 2019 - Oct. 2020

- In the school's logistics robot competition, I used Raspberry Pi to control the vehicle's movement and the gripping of the robotic arm to complete component assembly and debugging. In the end, we achieved the second prize in the university-level competition.

OTHERS

- I possess robust programming and implementation skills, specializing in Python and most of its libraries. I am proficient in the deep learning framework PyTorch and have considerable expertise in Linux.
- I have a competent command of English, as evidenced by an IELTS score of 6.5
- I enjoy playing basketball in my leisure time.