Youjia Zhang

Seoul, Korea | zhangyoujia@skku.edu | +82-010-5865-1115 | https://youjia-zhang.github.io

Research Interests

I am a fourth year Ph.D. student at SKKU (Sungkyunkwan University), South Korea, advised by (Prof. Sungeun Hong in the AI & Media Lab (AIM Lab). My recently research interests include VLM Pruning, Multimodal Learning, Audio-Visual Recognition, Parameter-Efficient Model Tuning and Test-Time Adaptation.

Education

Ph.D., Sungkyunkwan University, Seoul, South Korea — Mar 2024 – Present

- Major: Immersive Media Engineering
- Advisor: Prof. Sungeun Hong

Ph.D., Inha University, Seoul, South Korea — Sept 2021 – Mar 2024

- Major: Information and Communication Engineering
- Advisor: Prof. Sungeun Hong

M.S., Chongqing University of Posts and Telecommunications, Chongqing, China — Sept 2018 – June 2021

- Major: Computer Science and Technology
- Advisor: Prof. Xu Zhang

B.S., Chongqing University of Posts and Telecommunications, Chongqing, China — Sept 2014 – June 2018

- Major: Information and Computing Science

Projects

RGB-X Path Networks for Multi-modal Multi-task Learning — 2023.03 – 2026.02

- Funded by National Research Foundation of Korea (NRF)
- Developed path networks for RGB-X data (e.g., depth, thermal, tactile, text) to enable efficient multi-modal fusion and knowledge transfer across tasks and environments, thereby supporting generalization and adaptability in complex real-world settings.

RGB-D Object Detection and Segmentation based on Multimodal Fusion — 2023.03 – 2023.10

- Funded by Samsung Electronics
- Developed an object detection and segmentation framework that effectively fuses depth information with RGB images, enabling robots to automatically identify and manipulate specific objects.

Visuo-Tactile Perception for Human-Like Manipulation of Deformable Objects with Dynamic Center of Mass — 2021.09 – 2023.08

- Funded by Samsung Research Funding & Incubation Center for Future Technology
- Developed core technologies for stable grasping and manipulation of soft, deformable objects through "visual-tactile fusion" and "visual-tactile perception" for robots to manipulate objects at the human level.

Publications

Conference

1. Y. Zhang, Y. Kim, Y.G. Choi, H. Kim, H. Liu, and S. Hong, "Backpropagation-Free Test-Time Adaptation via Probabilistic Gaussian Alignment". Neural Information Processing Systems (NeurIPS) 2025.
2. Y. Cho*, H. Kim*, S. Kim, Y. Zhang, Y. Choi, and S. Hong, "RA-Touch: Retrieval-Augmented Touch Understanding with Enriched Visual Data". ACM Multimedia (MM) 2025.
3. H. Kim*, I. Jung*, D. Suh, Y. Zhang, S. Lee, and S. Hong, "Question-Aware Gaussian Experts for

Audio-Visual Question Answering". IEEE/CVF conference on computer vision and pattern recognition (CVPR) 2025.

4. S. Choi, Y. Zhang, and S. Hong, "Intra-inter modal attention blocks for rgb-d semantic segmentation". International Conference on Multimedia Retrieval (ICMR) 2023.
5. Y. Zhang, S. Choi, and S. Hong, "Spatio-Channel Attention Blocks for Cross-modal Crowd Counting". Asian Conference on Computer Vision (ACCV) 2022.
6. X. Zhang, Y. Zhang, and Z. Zhang, "Multi-granularity recurrent attention graph neural network for few-shot learning". International Conference on Multimedia Modeling (MMM) 2021.

Journal

1. Y. Zhang, H. Liu, Y. Kim, and S. Hong. "CAT-TPT: Class-Agnostic Text-based Test-time Prompt Tuning for Vision-Language Models". International Journal of Computer Vision (2025).
2. Y. Zhang, S. Choi, and S. Hong, "Memory-efficient cross-modal attention for RGB-X segmentation and crowd counting". Pattern Recognition (2025).
3. X. Zhang, D. Huang, H. Li, Y. Zhang, Y. Xia, and J. Liu, "Self-training maximum classifier discrepancy for EEG emotion recognition". CAAI Transactions on Intelligence Technology (2023).
4. X. Zhang, Y. Zhang, Z. Zhang, and J. Liu, "Discriminative learning of imaginary data for few-shot classification". Neurocomputing (2022).

Honors & Awards

SKKU STEM Full Scholarship	2024 – 2026
IEIE Encouraging Paper Award	2024
BK21 Excellent Research Award, Inha University	20023
ACCV Oral Presentation	2022
Global Vision Scholarship, Inha University	2024 – 2026
Mathematical Contest in Modeling(MCM), Meritorious Winners	2017

Patents

Spatio-channel attention blocks for cross-modal crowd counting (Registration number C-2022-055027)

Teaching Experience

TA, Advanced Computer Vision, Graduate Course, Sungkyunkwan University, Fall 2025

Lecturer & TA, Introduction to Deep Learning, Undergraduate Course, Sungkyunkwan University, Spring 2025

Lecturer & TA, Advanced Computer Vision, Graduate Course, Sungkyunkwan University, Fall 2024

TA, Computer Vision, Graduate Course, Inha University, Fall 2021

Academic Activities

Reviewer

- ACM International Conference on Multimedia (ACM MM)
- Neural Information Processing Systems (NeurIPS)
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)