

Kaining Ying

College of Computer Science and Artificial Intelligence
Fudan University

kaining.ying.cv@gmail.com
knying24@m.fudan.edu.cn
+86 17857345007
GitHub, Google Scholar

EDUCATION

- Ph.D. College of Computer Science and Artificial Intelligence, Fudan University, 2024–
Supervised by Prof. [Henghui Ding](#)
Research Topic: Multimodal Visual Segmentation
- M.E. College of Computer Science and Technology, Zhejiang University of Technology, 2021–2024
Supervised by Prof. [Zenhua Wang](#)
Research Topic: Instance Segmentation
- B.E. College of Computer Science and Technology, Zhejiang University of Technology, 2017–2021
Supervised by Prof. [Zenhua Wang](#)
Research Topic: Instance Segmentation, [Thesis](#)

RESEARCH EXPERIENCES

- 2023–2024 OpenGVLab, Shanghai Artificial Intelligence Laboratory
Research Intern, work with Dr. [Kaipeng Zhang](#), Dr. [Wenqi Shao](#) and Prof. [Ping Luo](#)
Research Topic: Evaluation of Large Vision-Language Models
- 2022–2023 State Key Lab of CAD & CG, Zhejiang University
Visiting Student, supervised by Prof. [Chunhua Shen](#) and Prof. [Hao Chen](#)
Research Topic: Video Instance Segmentation

RESEARCH AREAS

Multimodal LLM Evaluation

Segmentation: instance segmentation and video segmentation

PUBLICATIONS

* and [†] indicate equal contributions and corresponding authors respectively.

First-authored: 3 × ICCV, 1 × ICML.

- ICCV’25** K.Ying*, H. Hu*, H. Ding[†] (2025), “[MOVE: Motion-Guided Few-Shot Video Object Segmentation](#)”, In: *Proc. Int. Conf. Computer Vision*.
- ICCV’25** K.Ying, H. Ding[†], G. Jie, Y.G. Jiang (2025), “[Towards Omnimodal Expressions and Reasoning in Referring Audio-Visual Segmentation](#)”, In: *Proc. Int. Conf. Computer Vision*.
- Submission H. Ding, C. Liu, S. He, K. Ying, X. Jiang, C.C. Loy, Y.G. Jiang (2024), “[MeViS: A Multi-Modal Dataset for Referring Motion Expression Video Segmentation](#)”, *TPAMI Under Reviewed*.

- NeurIPS'24 S. Liu, K. Ying, H. Zhang, Y. Yang, Y. Lin, T. Zhang, C. Li, Y. Qiao, P. Luo, W. Shao, K. Zhang (2024), "[ConvBench: A Multi-Turn Conversation Evaluation Benchmark with Hierarchical Capability for Large Vision-Language Models](#)", In: *Proc. Advances in Neural Information Processing Systems*, **Spotlight**.
- ICML'24 K. Ying*, F. Meng*, J. Wang*, Z. Li, H. Lin, Y. Yang, H. Zhang, W. Zhang, Y. Lin, S. Liu, J. Lei, Q. Lu, P. Gao, R. Chen, P. Xu, R. Zhang, H. Zhang, Y. Wang, Y. Qiao, P. Luo, K. Zhang, W. Shao (2024), "[MMT-Bench: A Comprehensive Multimodal Benchmark for Evaluating Large Vision-Language Models Towards Multitask AGI](#)", In: *Proc. Int. Conf. Machine Learning*.
- ICONIP'23 Z. Wang, K. Ying, J. Meng, J. Ning (2023), "[Human-to-Human Interaction Detection](#)", In: *Proc. Int. Conf. on Neural Information Processing*, **Oral**.
- ICCV'23 K. Ying*, Q. Zhong*, W. Mao, Z. Wang[†], H. Chen[†], L.Y. Wu, Y. Liu, C. Fan, Y. Zhuge, C. Shen (2023), "[CTVIS: Consistent Training for Online Video Instance Segmentation](#)", In: *Proc. Int. Conf. Computer Vision*.
- TPAMI'23 J. Meng, Z. Wang[†], K. Ying, J. Zhang, D. Guo, Z. Zhang, Q. Shi, S. Chen (2023), "[Human Interaction Understanding with Consistency-Aware Learning](#)", *IEEE Trans. Pattern Analysis and Machine Intelligence*.
- TNNLS'22 P. Zhou*, K. Ying*, Z. Wang, D. Guo, C. Bai[†], (2022) "[Self-supervised Enhancement for Named Entity Disambiguation via Multimodal Graph Convolution](#)", *IEEE Trans. on Neural Networks and Learning Systems*.
- ICASSP'22 K. Ying, Z. Wang, C. Bai, P. Zhou (2022), "[ISDA: Position-Aware Instance Segmentation with Deformable Attention](#)", In: *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing*, **Oral**.

AWARDS AND HONORS

- | | |
|------|--|
| 2023 | 2nd Place in The 5th Large-scale Video Object Segmentation Challenge - Track 2: Video Instance Segmentation at ICCV 2023, Solution , Leaderboard , Certificate |
| 2023 | 2nd Place in Pixel-level Video Understanding Challenge (VPS Track) at CVPR 2023, Solution , Leaderboard , Certificate |
| 2022 | Distinguished Graduate Student of the College |
| 2022 | National Award Scholarship (Top 1) |

SERVICE

Peer Review

Journal: TIP, TMM, PR, Machine Vision and Applications, IET Computer Vision
 Conference: ICME (2023), ACM MM (2023, 2024, 2025), AAAI (2024, 2025), CVPR (2024), ICCV (2025), NeurIPS (2024, 2025), AISTATS (2025)

Updated June 2025