Mingqiao Ye

PhD Student in Computer Science · EPFL VILAB

mingqiao.ye@epfl.ch ◇ ymq2017.github.io ◇ GitHub ◇ Google Scholar

Research

My research **scales one model across many modalities and tasks**, rather than building a separate specialized model for each.

MODUS, my recent work, is a single decoder that generates any modality from any other (RGB, depth, segmentation, text, and learned features alike), with no modality-specific heads, losses, or pipelines. What interests me is what a single unified model makes possible: modalities that strengthen each other when learned together, and the same model fluent across perception, generation, and reasoning. Earlier, I built widely-used systems for high-quality and open-world visual understanding, including HQ-SAM, Gaussian Grouping, Cascade-DETR, and EntitySAM.

Education

EPFL, École Polytechnique Fédérale de Lausanne — Lausanne, Switzerland
Ph.D. in Computer Science, advised by Prof. Amir Zamir (VILAB) — Sep 2024 – Present

ETH Zurich — Zurich, Switzerland
M.Sc. in Electrical Engineering and Information Technology — Sep 2021 – May 2024

Zhejiang University — Hangzhou, China
B.Sc. in Electronic Information Engineering — Sep 2017 – Jun 2021

Experience

Adobe Research — San Jose, USA
Research Intern, advised by Dr. Joon-Young Lee — May 2024 – Aug 2024

Computer Vision Lab, ETH Zurich — Zurich, Switzerland
Research Assistant with Dr. Lei Ke and Dr. Martin Danelljan — 2022 – 2024

Publication Summary

1,482 citations ◇ h-index **5** ◇ **5,495** GitHub stars across released open-source projects.
Sources: Google Scholar and GitHub, as of June 2026.

Publications

* *denotes equal contribution.*

1. **MODUS: Decoder-only Any-to-Any Modeling of Diverse Modalities**
   Mingqiao Ye*, Zhaochong An*, Zhitong Gao, Xian Liu, Oğuzhan Fatih Kar, Jesse Allardice, Roman Bachmann, David Mizrahi, François Fleuret, Chuan Li, Amir Zadeh, Serge Belongie, Afshin Dehghan, Amir Zamir.
   ICML 2026 [Paper] [Project Page]

2. **EntitySAM: Segment Everything in Video**
   Mingqiao Ye, Seoung Wug Oh, Lei Ke, Joon-Young Lee.
   CVPR 2025 [Paper] [Code] ★65

3. **Stable Segment Anything Model**
   Qi Fan, Xin Tao, Lei Ke, Mingqiao Ye, Yuan Zhang, Pengfei Wan, Zhongyuan Wang, Yu-Wing Tai, Chi-Keung Tang.
   ICLR 2025 [arXiv] [Code] ★ 73

4. **Gaussian Grouping: Segment and Edit Anything in 3D Scenes**
   Mingqiao Ye, Martin Danelljan, Fisher Yu, Lei Ke.
   ECCV 2024 [arXiv] [Code] ★ 1.0k

5. **Segment Anything in High Quality (HQ-SAM)**
   Lei Ke*, Mingqiao Ye*, Martin Danelljan, Yifan Liu, Yu-Wing Tai, Chi-Keung Tang, Fisher Yu.
   NeurIPS 2023 [arXiv] [Code] ★ 4.2k

6. **Cascade-DETR: Delving into High-Quality Universal Object Detection**
   Mingqiao Ye*, Lei Ke*, Siyuan Li, Yu-Wing Tai, Chi-Keung Tang, Martin Danelljan, Fisher Yu.
   ICCV 2023 [arXiv] [Code] ★ 100

Awards & Honors

- Swiss AI Initiative: 2025 Fellowship Program.
- EPFL Computer and Communication Sciences Doctoral Program (EDIC): Fellowship for the first year of Ph.D. studies, 2024.
- M.Sc. degree awarded with distinction (*mit Auszeichnung*), ETH Zurich, 2024.

Teaching

- Teaching Assistant, COM-304 Intelligent Systems: Communications & AI, EPFL (2025, 2026).
- Teaching Assistant, CS-503 Visual Intelligence, EPFL (2025).
- Teaching Assistant, CS-119(d) Information, Computation, Communication, EPFL (2025).
- Main contributor to nanoMFM, an educational codebase with minimal implementations of GPT, MaskGIT, 4M, Flow Matching, and a VLM, used in COM-304 and CS-503.

Academic Service

Conference Reviewer
CVPR (2025, 2026), ICLR (2025, 2026), ICML (2026), NeurIPS (2026), ECCV (2026), ACM MM (2026), SIGGRAPH Asia (2026), BMVC (2026).

Journal Reviewer
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), International Journal of Computer Vision (IJCV), IEEE Transactions on Geoscience and Remote Sensing (TGRS), Computational Visual Media (CVMJ), Machine Intelligence Research (MIR).

Workshop Reviewer
ICCV 2025 Workshop on Generating Digital Twins from Images and Videos (gDT-IV); Any-to-Any Multimodal Learning (A2A-MML), CVPR 2026.

Skills

Programming	Python, C++, CUDA
Frameworks & Tools	PyTorch, Hugging Face, distributed / multi-GPU training, Git, Docker, Linux, Slurm, LaTeX
Languages	Mandarin (native), English (fluent)