Binjie Zhang

Google Scholar

Nationality: Chinese
Mobile: +65 80273505 / +86 15216801797
Email: binjie97@u.nus.edu / zbj19@tsinghua.org.cn

Education

- **National University of Singapore (NUS)** — Singapore
  Ph.D. candidate in Computer Science — *Jan 2023 – Jan 2027 (expected)*
  Supervisor: *Assistant Prof. Mike Zheng Shou*
  Research Interests: *Video understanding, vision-language models, world models*

- **Tsinghua University (THU)** — Beijing, China
  M.Eng. in Computer Science and Technology — *Aug 2019 – Jul 2022*
  Supervisor: *Prof. Chun Yuan*
  Research Focus: *Compatible representation learning, cross-modal video understanding*

- **East China University of Science and Technology (ECUST)** — Shanghai, China
  B.Eng. in Information Engineering — *Aug 2015 – Jul 2019*
  Cumulative GPA: *3.77 / 4.00*, **Ranking:** *1 / 92*

Publications

Selected first-author works in top-tier conferences

- [1] "GR-MCT: Group-Relative Reasoning with Multi-Modal CoT for Tool-Using Agents." *CVPR* (under submission), 2026.
- [2] "PRCFC: Lifelong Imitation Learning via Prototype Replay and Coarse-to-Fine Compatibility." *CVPR* (under submission), 2026.
- [3] "Ego-centric Predictive Model Conditioned on Hand Trajectories." *ICLR* (under review), 2025.
- [4] "TaCA: Upgrading Your Visual Foundation Model with a Task-Agnostic Compatible Adapter." arXiv, 2023.
- [5] "Darwinian Model Upgrades: Model Evolving with Selective Compatibility." *AAAI*, 2023.
- [6] "Towards Universal Backward-Compatible Representation Learning." *IJCAI* (long oral), 2022.
- [7] "Hot-Refresh Model Upgrades with Regression-Alleviating Compatible Training in Image Retrieval." *ICLR*, 2022.

Research & Work Experience

- **AI Research Intern, Tencent ARC Lab** — Shenzhen, China
  Compatible Representation Learning — *2020 – 2023*
  - Researched backward-compatible representation learning and model upgrades for large-scale image/video retrieval systems.
  - Implemented and maintained training pipelines (multi-GPU, large-scale feature extraction) for CLIP-style and contrastive models.
  - Designed and evaluated methods to reduce model regression during hot-refresh upgrades, enabling new models to be deployed without re-encoding full galleries.

- **AI Research Intern, Tencent ARC Lab** — Shenzhen, China
  Cross-Modality Video Understanding — *2019 – 2020*
  - Developed multi-modal deep learning models for video-text retrieval and temporal grounding on large-scale datasets.
  - Built data preprocessing, training, and evaluation codebases in PyTorch and supported internal research experiments.

Honors and Awards

- Tencent Technology Breakthrough Award - Hot-Refresh Model Upgrades — 2022
- SZCCF Science and Technology Award - Efficient Model Upgrades — 2022
- Excellent Master Degree Graduate in Beijing & Outstanding Master's Graduation Thesis — 2022
- Annual College Personage Award (highest student honor in ECUST) — 2018
- National Scholarship for Undergraduates (twice), Ministry of Education of China — 2017, 2016

Selected Projects

- **Robot Lifelong Learning**[2]: Developed PRCFC, a lightweight framework for lifelong imitation learning using compact prototype replay and compatibility regularization, reducing forgetting and improving cross-task transfer on LIBERO.
- **Ego-Centric Predictive Model**[3]: Designed a two-stage model that predicts future hand trajectories and uses them to guide a Latent Diffusion Model for egocentric future video generation, achieving state-of-the-art results on Ego4D, BridgeData, and RLBench.
- **Task-Agnostic Compatible Adapter**[4]: Proposed TaCA, a parameter-efficient adapter that enables seamless upgrades between visual foundation models (e.g., CLIP variants) without retraining downstream tasks, validated on large-scale video-language benchmarks.
- **Hot-Refresh Model Upgrades**[7]: Studied model regression in hot-refresh upgrades for large-scale image retrieval systems, proposing regularization and uncertainty-based backfilling strategies to maintain compatibility without re-encoding existing galleries.