Yichao Cai

yichao.cai@adelaide.edu.au · +61 478 927 202 · yichaocai.com · GitHub · Google Scholar · LinkedIn

Research Statement

I study how language supervision shapes the semantics, geometry, and identifiability of multimodal representations. My long-term goal is to develop principled foundations for language-guided representation learning, with the aim of building multimodal systems that acquire structured, interpretable, and reliable representations of the world.

Education

Ph.D. in Computer Science *Sep 2023 – Present*
Adelaide University (formerly The University of Adelaide)
Supervisor: Prof. Javen Qinfeng Shi

M.Sc. in Instrument Science and Technology *Sep 2016 – Jun 2019*
Wuhan University of Technology
Supervisor: Prof. Xiao Zhou

B.Eng. in Measurement & Control Technology and Instrument *Sep 2012 – Jun 2016*
Wuhan University of Technology

Experience

PhD Student Researcher *Sep 2023 – Present*
Australian Institute for Machine Learning (AIML), Adelaide University

- Developed a theoretical framework for the geometry of InfoNCE-based contrastive learning, characterising alignment potentials, entropic dispersion forces, and the emergence of modality gap (ICMLR 2026).
- Studied cross-modal misalignment as a beneficial inductive bias in multimodal contrastive learning; established identifiability conditions and controlled benchmarks (NeurIPS 2025 Spotlight).
- Developed CLAP, a contrastive representation learning framework for content–style disentanglement via augmented language prompts (ECCV 2024).

AI Engineer *Jun 2020 – Aug 2022*
Tellhow Software

- Built inspection and safety-analytics solutions for power-industry equipment using computer vision and automated analysis workflows; delivered end-to-end features from prototyping to production deployment.

Software Engineer *Jul 2019 – Apr 2020*
Huawei Technologies

- Developed and maintained communication service components; implemented production features and system integrations.

Visiting Student Researcher *May 2018 – Sep 2018*
California PATH, UC Berkeley

- Conducted research on visual perception algorithms for autonomous driving.

Research Interests

- **Representation learning:** learning objectives and training paradigms; identifiability theory; semantic structure in learned representations.
- **Vision-language modeling:** multimodal alignment; multimodal large language models; supervision design and data curation for multimodal learning.
- **Explainable machine learning:** mechanistic interpretability; representation geometry; latent-structure characterization.

Teaching

Teaching Assistant - Neural Networks and Deep Learning (ARTI X300), Adelaide University *Semester 1 2026*
Guest Lecturer and Head Tutor - Statistical Machine Learning (COMP SCI 3314), Adelaide University *Semester 2 2025*
Teaching Assistant - Using Machine Learning Tools (COMP SCI 7317), Adelaide University *Trimester 2 2025*
Teaching Assistant - Concepts in AI and ML (COMP SCI 7327), Adelaide University *Semester 1 2025*

Academic Service

Conference Reviewers

- Conference on Neural Information Processing Systems (NeurIPS) 2026
- International Conference on Machine Learning (ICML) 2026, Silver Reviewer Award
- International Conference on Learning Representations (ICLR) 2026

Journal Reviewers

- Transactions on Machine Learning Research (TMLR)

Honors and Awards

NeurIPS 2025 Scholar Award, Neural Information Processing Foundation *Oct 2025*
University of Adelaide Research Scholarships *Sep 2023*
Award for Outstanding Graduates, Wuhan University of Technology *Jun 2019*

Publications

[1] **Yichao Cai**, Zhen Zhang, Yuhang Liu, Javen Qinfeng Shi. **The Geometric Mechanics of Contrastive Representation Learning: Alignment Potentials, Entropic Dispersion, and Modality Gap.** *International Conference on Machine Learning (ICML), 2026*.

[2] Wenkang Jiang, Yuhang Liu, **Yichao Cai**, Erdun Gao, Jiayi Dong, Ehsan Abbasnejad, Lina Yao, Javen Qinfeng Shi. **What Makes a Good Representation for Single-Cell Perturbation Prediction?** *International Conference on Machine Learning (ICML), 2026*.

[3] Jiaqing Chen, Zidu Yin, **Yichao Cai**, Yuhang Liu, Zhen Zhang, Dong Gong, Javen Qinfeng Shi. **Boundary Embedding Shaping with Adaptive Contrastive Learning for Graph Structural Disentanglement.** *International Conference on Machine Learning (ICML), 2026.*

[4] Yuhang Liu, Dong Gong, **Yichao Cai**, Erdun Gao, Zhen Zhang, Biwei Huang, Mingming Gong, Anton van den Hengel, Javen Qinfeng Shi. **I Predict Therefore I Am: Is Next Token Prediction Enough to Learn Human-Interpretable Concepts from Data?** *International Conference on Learning Representations (ICLR), 2026*.

[5] **Yichao Cai**[*], Yuhang Liu[*], Erdun Gao, Tianjiao Jiang, **Zhen Zhang, Anton van den Hengel, Javen Qinfeng Shi. On the Value of Cross-Modal Misalignment in Multimodal Representation Learning.** *Advances in Neural Information Processing Systems (NeurIPS), 2025 (Spotlight).* [*]Equal contribution.

[6] **Yichao Cai**, Yuhang Liu, Zhen Zhang, Javen Qinfeng Shi. **CLAP: Isolating Content from Style Through Contrastive Learning with Augmented Prompts.** *European Conference on Computer Vision (ECCV), 2024*.

Miscellaneous

Reading, poetry writing, guitar, and nature walks.