Shuanghao Bai

Tel: +86-199-4940-8829 | Email: baishuanghao@stu.xjtu.edu.cn

Homepage | Google Scholar | Github

Xi'an, Shannxi - China

Education

- **Xi'an Jiaotong University** *Sept. 2022 - Jun. 2027*
  Ph.D. candidate. of Control Science and Technology Xi'an, China
  - Advisor: Badong Chen
  - Research Interests: Generalization of Machine Learning, Robotics.
- **Westlake University** *Sept. 2024 - Feb. 2025*
  Visiting Student in MiLAB Hangzhou, China
  - Advisor: Donglin Wang
  - Research Interests: Robotics.
- **Chongqing University** *Sept. 2018 - Jun. 2022*
  Bachelor of Automation Chongqing, China
  - Advisor: Min Zhao
  - GPA: 3.68/4.00 (Top 5%)

Projects

- **Cloud-Edge-Device Robot Platform** *Sept. 2022 - Dec. 2025*
  Basic theories and key technologies of cloud-edge-device integrated service robot cloud-brain platform
  - Mainly focuses on generalization tasks in computer vision, addressing the challenges posed by limited data availability and significant distribution shifts between training and test data.
  - Extend the utility of multilayer perceptron to cross-domain few-shot classification. [GitHub]
  - Implement prompt tuning vision-language model CLIP to unsupervised domain adaptation. [GitHub]
  - Applied prompt tuning vision-language model CLIP to domain generalization. [GitHub]
- **Robotic Arm Platform (Project Applicant and Leader: Shuanghao Bai)** *Jan. 2024 - Dec. 2025*
  Robotic arm platform technology and application based on visual language action model
  - Developed a vision language action model that enables robot manipulation.
- **Multi-agent Collaboration** *Jun. 2023 - Dec. 2023*
  Research on natural human-machine interaction technology for heterogeneous unmanned swarms
  - The system primarily enables robots (drones and little car fleets) to understand human language and make decisions based on environmental perception.
  - Language comprehension involves task decomposition and code generation using a large language model. Environmental perception relies on drones capturing RGB images. A vision-language model analyzes these images to generate heatmaps, which guide the cars' actions.

Honors and Awards

- **National Scholarship** *Dec. 2024*
- **National Third Prize in the Phoenix Intelligent Technology Innovation and Application Competition** *Jun. 2021*
- **Grade A Comprehensive Scholarship in Chongqing University** *Dec. 2020*
- **Outstanding Individual Youth Volunteer of Chongqing University** *May. 2020*
- **National Scholarship** *Dec. 2019*
- **Outstanding Student of Chongqing University** *Dec. 2019*

SKILLS

- **Programming Languages:** Python, Pytorch, C++
- **Languages:** Chinese, English

PUBLICATIONS

C=CONFERENCE, J=JOURNAL, S=IN SUBMISSION

I. Generalization in Vision Language Models

[C.1] **Shuanghao Bai***, Yuedi Zhang*, Wanqi Zhou, Zhirong Luan, Badong Chen. Soft Prompt Generation for Domain Generalization. In European Conference on Computer Vision (ECCV). 2024. [Paper] [Code]

[C.2] **Shuanghao Bai**, Min Zhang, Wanqi Zhou, Siteng Huang, Zhirong Luan, Donglin Wang, Badong Chen. Prompt-based Distribution Alignment for Unsupervised Domain Adaptation. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). 2024. [Paper] [Code]

[C.3] **Shuanghao Bai**, Wanqi Zhou, Zhirong Luan, Donglin Wang, Badong Chen. Improving Cross-domain Few-shot Classification with Multilayer Perceptron. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2024. [Paper] [Code]

[C.4] Haoran Zhang*, **Shuanghao Bai***, Wanqi Zhou, Jingwen Fu, Badong Chen. PromptTA: Prompt-driven Text Adapter for Source-free Domain Generalization. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2025. [Paper] [Code]

[S.1] Wanqi Zhou, **Shuanghao Bai**, Qibin Zhao, Badong Chen. Revisiting the Adversarial Robustness of Vision Language Models: a Multimodal Perspective. ArXiv preprint arXiv: 2404.19287. [Paper] [Code]

II. Robot Learning

[C.1] Wei Zhao, Pengxiang Ding, Zhang Min, Zhefei Gong, **Shuanghao Bai**, Han Zhao, Donglin Wang. VLAS: Vision-Language-Action Model with Speech Instructions for Customized Robot Manipulation. In International Conference on Learning Representations (ICLR). 2025. [Paper] [Code]

[J.1] Zhirong Luan, Yijun Lai, Rundong Huang, **Shuanghao Bai**, Yuedi Zhang, Haoran Zhang, Qian Wang. Enhancing Robot Task Planning and Execution through Multi-Layer Large Language Models. In Sensors. 2024. [Paper]

[S.1] **Shuanghai Bai**, Wanqi Zhou, Pengxiang Ding, Wei Zhao, Donglin Wang, Badong Chen. Rethinking Latent Representations in Behavior Cloning: An Information Bottleneck Approach for Robot Manipulation. ArXiv preprint arXiv: 2502.02853 [Paper] [Project] [Code]

III. Causal Learning in Machine Learning

[C.1] Wanqi Zhou, **Shuanghao Bai**, Shujian Yu, Qibin Zhao, Badong Chen. Jacobian Regularizer-based Neural Granger Causality. In International Conference on Machine Learning (ICML). 2024. [Paper] [Code]

[S.1] Wanqi Zhou, **Shuanghao Bai**, Yicong He, Badong Chen. An Information-Theoretic Approach for Heterogeneous Differentiable Causal Discovery. In Nerual Networks. 2025. [Paper] [Code]

ACADEMIC SERVICE

- **Conference Reviewer:** ICIRA
- **Journal Reviewer:** TIP, TCSVT, KBS, Neural Networks, Neurocomputing

ABOUT ME

As a third-year direct Ph.D. candidate at Xi'an Jiaotong University, I'm deeply fascinated by computer vision, with a particular focus on generalization in computer vision and its applications in robotics. The more I learn, the more I realize how much there is to explore in these fields!

I am actively seeking academic and industrial exchange opportunities for Fall 2025, specifically focusing on joint Ph.D. programs and internship projects. My hope is to find a research team where I can roll up my sleeves, dive into some cutting-edge projects, and both contribute my skills and learn new ones. I'm eager to experience a different academic environment and see how it shapes my perspective on research.