ZHOU BAICHUAN

Shanghai AI Lab, Shanghai
zhoubaichuan@pjlab.org.cn ◇ Website ◇ GitHub ◇ Google Scholar

EDUCATION

B.Eng, Beihang University 2020 - 2024
School of Automation Science and Electrical Engineering
Grades: 88/100
English Proficiency: **IELTS 8.0**, CET6 663, CET4 667

PUBLICATION & PREPRINTS

1. (ICLR 2025 under review) LOKI: A Comprehensive Synthetic Data Detection Benchmark using Large Mnulti-modal Models. (Project Page)
   Junyan Ye[*], ***Baichuan Zhou***[*], *Zilong Huang*[*], *Junan Zhang*[*], *Tianyi Bai, Hengrui Kang , Jun He, Honglin Lin, Zihao Wang, Tong Wu, Zhizheng Wu, Yiping Chen, Dahua Lin, Conghui He, Weijia Li*
2. (AAAI 2025 under review) UrBench: A Comprehensive Benchmark for Evaluating Large Multimodal Models in Multi-View Urban Scenarios. (Project Page)
   Baichuan Zhou[*], *Haote Yang*[*], *Dairong Chen*[*], *Junyan Ye*[*], *Tianyi Bai, Jinhua Yu, Songyang Zhang, Dahua Lin, Conghui He, Weijia Li*
3. (Arxiv Preprint) TinyLLaVA: A Framework of Small-scale Large Multimodal Models. (Project Page)
   Baichuan Zhou, *Ying Hu, Xi Weng, Junlong Jia, Jie Luo, Xien Liu, Ji Wu, Lei Huang*

[*]These authors contributed equally to this work.

OPEN SOURCE CONTRIBUTIONS

- **TinyLLaVA**. A **framework** for **training, deploying, benchmarking and visualizing** small-scale large multimodal models. **Project Leader**, **600+ Stars**, **350K+ checkpoint downloads**.
- **LOKI**. An **evaluation framework** for **multimodal synthetic detection** that supports **20+ mainstream foundation models** across audio, 3D, text, image, video modalities. **Core contributor**. **80+ stars**.

RESEARCH EXPERIENCE

Research Intern @ OpenDataLab / Shanghai AI Lab 2024.04 - Now
Advised by Prof. Weijia Li and Dr. Conghui He

1. **Multimodal Synthetic Data Detection**
   - **Implemented** a **multimodal evaluation framework** that **unifies** APIs of **20+ mainstream foundation models** across **5 modalities** with **data parallel** and **tensor parallel** support.
   - **Evaluated** and **documented 20+** multimodal foundation **models** and **provided analysis** and **results** for the performance of different models.
2. **Probing Foundation Models as Visual Assistants under Urban Environment**
   - **Major role** in curating a **comprehensive multi-view urban benchmark** for **foundation models**.
   - **Setup codebase** and **evaluated 20+ LMMs** in the **urban environment** and **analyzed** their performance.

Research Assistant @ 42 / Beihang University 2023.04 - 2024.04
Advised by Prof. Lei Huang

1. **Efficient Training and Inference of Large Multimodal Models**
   - **Setup codebase** and **datasets** and **trained the TinyLLaVA model family** that consistently **outperform bigger counterparts**, i.e., LLaVA-1.5, with **limited compute** and **data**.
   - **Babysat** and **run most** of the **ablation studies**.

SKILLS

Programming Languages	Python(proficient), C/C++/MATLAB/Java(experience)
Deep Learning Frameworks	PyTorch, Transformers, DeepSpeed
MISC	Javascript, React, TensorFlow.js

HONORS

2021 SMC Scholarship Fall, 2021

Beihang University Academic Competition Scholarship 2021, 2022, 2023