Jingkun An

anjingkun02@gmail.com · (+1) 919-713-8066 · Google Scholar · Homepage

Education

Beihang University, Beijing August 2020 – June 2024

Bachelor of Engineering, Major in Software Engineering, GPA 3.78/4.00.

Technical Skills

- **Programming Languages**: Python, Java, C
- **Artificial Intelligence**: Familiar with machine learning and deep learning fundamentals, Diffusers, Transformers, LLM application and fine-tuning, proficient in Scikit-Learn, PyTorch, etc.
- **Application Development**: Experienced with Django; familiar with Docker, MySQL, etc.

First Author Paper

3D Spatial Reasoning Model with Advanced Dataset and Benchmark Dec 2024 – May 2025

- **Publication:** "RoboRefer: Towards Spatial Referring with Reasoning in Vision-Language Models for Robotics", **accepted** in *Conference on Neural Information Processing Systems (NeurIPS)*, 2025.
- **Description:** We developed a novel 3D-aware Vision-Language Model for multi-step spatial referring with explicit reasoning. To train and evaluate the *RoboRefer* model, **I not only defined 31 fine-grained spatial relations to construct *RefSpatial*, the largest spatial understand and referring dataset, which contains over 20 million high-quality QA pairs and human-like reasoning processes, but also built *RefSpatial-Bench*, a complex multi-step spatial referring benchmark.** Achieving a new SOTA, the *RoboRefer* model gets an average accuracy of 89.6% on spatial relation understanding benchmarks and surpassed *Gemini 2.5 Pro* by 17.4% on complex spatial referring tasks. **My open-sourced dataset and benchmark have been downloaded over 15,000 times.** *RefSpatial-Bench* has been used by *Qwen3-VL* and *Gemini-Robotics-1.5* to evaluate their capabilities of spatial referring in complex scenes. [Project page].
- **Supervisor:** Lu Sheng (Prof. of Beihang University), Shanghang Zhang (Prof. of Peking University)

3D trace planning Model with Advanced Dataset and Benchmark Aug 2025 – Nov 2025

- **Publication:** "RoboTracer: Mastering Spatial Trace with Reasoning in Vision-Language Models for Robotics", **under review** at *Computer Vision and Pattern Recognition (CVPR)* 2026.
- **Description:** we proposed *RoboTracer*, a 3D-aware VLM that first achieves both 3D spatial referring and measuring as key perceptual cues for accurate spatial trace generation. Based on my expertise with 3D dataset from "*RoboRefer*" project, **I led the team to construct *TraceSpatial*, a large-scale dataset of 30M QA pairs, covering outdoor/indoor/tabletop scenarios, to enable our model "*RoboTracer*" to learn 3D spatial referring and measuring in SFT, and further achieve spatial tracing in RFT.** On spatial measuring benchmarks, *RoboTracer* achieves SOTA success rate of 76.3%, exceeding Gemini-2.5 Pro by 14%. To address the lack of benchmarks for spatial tracing, **I built *TraceSpatial-Bench*, which contains 100 real-world images with manually annotated tasks involving 3D object localization/movement/placement.** *RoboTracer* still achieves best performance on this benchmark, surpassing *Gemini-2.5 Pro* by 36%.
- **Supervisor:** Lu Sheng (Prof. of Beihang University), Shanghang Zhang (Prof. of Peking University)

Enhancing Text-to-Image Generation via Fully AI-Driven Feedback Dec 2023 – Aug 2024

- **Publication:** "AGFSync: Leveraging AI-Generated Feedback for Preference Optimization in Text-to-Image Generation", **accepted** in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025.
- **Description:** **I developed a fully AI-driven framework to enhance the performance of T2I models (e.g., SDXL and SD v1.5) through Direct Preference Optimization (DPO).** The approach first employs a self-training methodology through modeling constrains derived from instructions and the physical world to autonomously generate preference data, which is then used to fine-tune the model via DPO. This low-cost process, requiring no human annotation, improves SDXL's text consistency and aesthetic properties by 7% and generated higher-quality images in 67% of samples compared to prior methods. Ultimately, the method surpasses the previous SFT dataset approach (MJHQ-30K) by 5%. [Project page].
- **Supervisor:** Yemin Shi (PhD, Peking University; Co-founder & CTO of Dynamics Lab)

Bias Mining and Mitigation Algorithm Dec 2022 – Jun 2023

- **Publication:** "M3Fair: Mitigating Bias in Healthcare Data through Multi-Level and Multi-Sensitive-Attribute Reweighting Method", abstract **accepted** for poster in *Health Data Science, 2023, Beijing*.
- **Description:** **I developed a custom algorithm to detect and address biases across multiple features in medical electronic health records.** The core of the method involves classifying multiple sensitive-attribute features into different bias levels and then performing a multi-level reweighting operation. This approach is both generalizable and scalable, allowing for the flexible customization of fairness levels within clinical decision-making processes. In practice, this approach proved highly effective, successfully mitigating bias in 91.67% of samples and surpassing the performance of previous mitigation methods on 83.3% of samples. This algorithm won the **third prize** (Top 5) in 2023 NIH Bias Detection Challenge. [Project page].
- **Supervisor:** Lu An (PhD, NC State University; AI Engineering Manager of NVIDIA)

Internship & Research Experience

School of Computer Science, Peking University Dec 2024 – Present

- **Description:** During this internship, I took charge of the 3D spatial dataset construction of the *RoboRefer* and *RoboTracer* projects. Utilizing my *RefSpatial dataset*, **I assisted in training the RoboBrain 2.0 model for the research group, which, thanks to this foundational dataset, surpassed Gemini 2.5 Pro by 20% in spatial understanding.** My *RefSpatial-Bench* fills the gap in evaluating spatial referring with multi-step reasoning, while my *TraceSpatial-Bench* is the first 3D spatial tracing benchmark, which can evaluate models' capabilities of 3D trajectory planning.

School of Electronic and Computer Engineering, Peking University Dec 2023 – Aug 2024

- **Description:** I contributed to an interior design image generation model, **gaining experience in training diffusion models and applying alignment methods such as DPO and SFT**. Through a collaboration between the university lab and Shenzhen RabbitPre AI Technology Company, I applied *AGFSync* framework that I developed to improve the performance of a image generation model for commercial application.

Project Experience

TrustMe: A trustworthy AI platform to assist clinical decision-making Dec 2022 – Sep 2023

- **Description:** Built a trustworthy AI platform to assist clinical decision-making based on electronic medical records. The platform integrates data analysis with our custom fairness detection and mitigation algorithms, surpassing previous AI diagnostic models by 10%. Demonstrating its real-world value, **the system has been adopted by multiple top-tier Grade 3A hospitals in China.** [News]
- **Role:** Project Leader: Coordinated team members' work and developed the core bias detection algorithms.
- **Achievements:** This project won multiple national awards including the Grand Prize **(First Place)** at the 2023 "Challenge Cup" National Competition.

Awards

- 2023 National College Student "Challenge Cup" (National Grand Prize, Top 1)
- 2023 NIH Bias Detection Competition (International Third Prize, Top 5)
- 2023 "Internet+" Innovation and Entrepreneurship Competition (Second Prize, Top 10%)
- 2023 & 2024 Beijing College Student Entrepreneurship Project (Excellence Award, Top 5%)
- 2023 Fengru Cup Main and Industrial Track at Beihang University (Second Prize, Top 3%)
- 2024 Outstanding Graduate of Beijing (Top 5%)
- 2024 May Fourth Medal Nomination Award (Top 0.1%)
- 2023 Aviation Industry Scholarship (Top 1%)
- 2023 Star of the School of Software (Top 3%)
- 2023 Merit Student of Beihang University (Top 5%)

Additional Information

- **Languages**: TOEFL best Score 103