Kevin Zhai

+1 (407) 480-9635 • github.com/k-zhai • kevin.zhai@ucf.edu

Summary

- Ph.D. researcher in Computer Science at the University of Central Florida, advised by Dr. Mubarak Shah.
- Specialize in inference-time alignment, safety, and personalization for diffusion and diffusion-LLM models.
- Previous experience as a computer engineer building simulation systems; background in computer science, physics, and mathematics (triple major).

Research Interests

- Generative modeling: diffusion models, diffusion LLMs, flow matching.
- Inference-time alignment and test-time scaling.
- Safety, personalization, and unlearning in text-to-image models.

Selected Publications & Manuscripts

1. [**ACCEPTED**] Lee, J., Moon, H., **Zhai, K.**, et al. *Test-Time Scaling in Diffusion LLMs via Hidden Semi-Autoregressive Experts*. **Published at ICLR 2026**.
2. **Zhai, K.**, Singh, U., et al. *MIRA: Mitigating Reward Hacking in Inference-Time Noise Optimization for Text-to-Image Diffusion Models*. Submitted to ICML 2026.
3. **Zhai, K.**, Mollah, S., et al. *CORE: Context-Robust Remasking for Diffusion Language Models*. Submitted to ICML 2026.
4. **Zhai, K.**, Singh, U., et al. *Consensus-to-Personal: Inference-Time Alignment for Text-to-Image Personalization*. Submitted to CVPR 2026.
5. Ghosh, S., **Zhai, K.**, et al. *Repair-Aware Forgetting: An Iterative Approach to Unlearning in T2I Diffusion Models*. Submitted to ICML 2026.

Research and Work Experience

Graduate Research Assistant Aug 2023 – Present
Institute of Artificial Intelligence (IAI), University of Central Florida Orlando, FL

- Research on generative models, focusing on inference-time alignment, safety, and test-time scaling for diffusion and diffusion-LLMs.
- **First Author of MIRA (Submitted to ICML 2026):** Designed a training-free inference-time alignment algorithm that mitigates reward hacking via score-based KL regularization, outperforming standard guidance methods.
- **First Author of CORE (Submitted to ICML 2026):** Developed "Context-Robust Remasking," a framework for diffusion LLMs that identifies and corrects structural inconsistencies via targeted context perturbation.
- **First Author of COPE (Submitted to CVPR 2026):** Architected a two-stage "Consensus-to-Personal" framework using a single safety-aligned backbone with group-specific inference-time steering.
- **Co-Author of HEX (Accepted to ICLR 2026):** Contributed to a test-time scaling method for diffusion LLMs via heterogeneous block schedules; led ablation design and analysis of decoding optimization.
- Developed and maintained distributed experimental pipelines in Python/PyTorch on SLURM multi-GPU clusters, managing data preparation, job arrays, logging (WandB), and automated evaluation.

Computer Engineer Jul 2021 – Jun 2023
DCS Corp Alexandria, VA

- Developed algorithms to mimic human driving behavior in ground-vehicle simulations, combining physics-based models with behavior rules.

- Planned and built a testing environment in Unreal Engine 4 to approximate real-world driving conditions for evaluating autonomy stacks.

Undergraduate Researcher May 2020 – May 2021
Vanderbilt University Nashville, TN

- Developed network-simulation software to study fidelity under adversarial conditions.
- Explored approaches for mitigating network attacks using simulation-driven analysis.

Technical Skills

Languages:	Python, C++, JavaScript, Bash
ML / AI:	Diffusion models, diffusion LLMs, deep learning, RL, preference optimization
Frameworks:	PyTorch (DDP, FSDP), Hugging Face Diffusers, Accelerate, SLURM, WandB, Docker
Other:	High-performance computing, experiment design, data visualization

Education

Ph.D. in Computer Science Aug 2023 – Present
University of Central Florida, Center for Research in Computer Vision Orlando, FL

B.Sc. in Computer Science, Physics, and Mathematics (Triple Major) Aug 2017 – May 2021
Vanderbilt University Nashville, TN

Teaching and Mentoring

Graduate Teaching Assistant – Data Structures Fall 2024 – Spring 2025
University of Central Florida

- Helped students learn core data structures and debugging strategies through labs and office hours.
- Contributed to a course pass rate above 80% (historically around 50%) by providing targeted support on programming assignments.
- Gave individual feedback on code quality and problem-solving approaches.

Grants and Awards

- UCF ORCGS Doctoral Fellowship ($25,000 per year)