ZEKAI YIN | zekaiyin2025phd@gmail.com | San Jose, CA

WORK EXPERIENCE

Black Sesame Technologies, San Jose, CA **Research Scientist, Deep Learning Algorithm** | Aug 2025 – Present

- Research-engineered an end-to-end, **multimodal Flow-Matching trajectory planning** model that encodes trajectories into discrete tokens using VQ-VAE and RFSQ, improving ADE by **20%** while enabling real-time inference with only 2 denoising steps, optimizing hyperparameters through Bayesian optimization techniques.
- Designed an anchor-based trajectory scoring model that leverages **diffusion likelihood** estimates to **rank multimodal** predictions, improving selection of the safest and most optimal path across diverse future anchors in complex driving scenarios.
- Led the development of a **multimodal**, goal-based agent prediction module using a heatmap-offset formulation, improving precision and recall for detecting cut-in, crossing, and intrusion scenarios by **35%.**
- Developed a navigation-aware data augmentation pipeline for lane-change training, improving navigation-conditioned lane-change success rate from 50% to **90%**.
- Optimized the model for **embedded deployment** using **ONNX**, ensuring strict real-time constraints for autonomous driving.

EDUCATION

Boston University, Boston, USA | Sep 2023 - Jan 2025 **Master of Science in Artificial Intelligence**
Coursework: Deep Learning, Computer Vision, Computer Systems, Neural Networks, Reinforcement Learning, MLOps
Peking University, Beijing, China | Sep 2019 - Jul 2023 **Bachelor of Science in Data Science and Big Data Technology**
Coursework: Statistical Learning, Data Mining, Natural Language Processing, Distributed Systems, Operating Systems

PUBLICATIONS (* for equal contribution)

Robot Structure Prior Guided Temporal Attention for Camera-to-Robot Pose Estimation from Image Sequence
Zekai Yin*, *Yang Tian*, Jiyao Zhang*, Hao Dong*
Accepted *by IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), **2023***

ZeroVO: Visual Odometry with Minimal Assumptions
Zekai Yin*, *Lei Lai*, Eshed Ohn-Bar*
Accepted *by IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), **2025***

BranchOut: Capturing Realistic Multimodality in Autonomous Driving Decisions
*Hee Jae Kim**, Zekai Yin,** Lei Lai, Jason Lee, Eshed Ohn-Bar*
Accepted *by The Conference on Robot Learning (**CoRL**), **2025***

DOGMA: A Wearable Multimodal Dog-Human Behavior Dataset in the Wild (Under Review)
*Tianyou Luo, Adwait Kulkarni, Hee Jae Kim, **Zekai Yin**, Maolin Wei, Rohit Mandapati, Reza Sajjadinasab, Eshed Ohn-Bar*

GuidedWays: A Naturalistic Navigation Dataset of Blind People and Their Guide Dogs (Under Review)
*Hee Jae Kim, Maolin Wei, Tianyou Luo, Reza Sajjadinasab, **Zekai Yin**, Hernisa Kacorri, Eshed Ohn-Bar*

RESEARCH EXPERIENCE

H2X Lab, Boston University **Machine Learning Engineer / Research Assistant** | Jan 2024 – Aug 2025

- Proposed **ZeroVO** (**Zero**-Shot **V**isual **O**dometry) as a co-first author, presenting a novel visual odometry algorithm by fusing **VLM** features with geometric information using a multi-head-cross-attention module, accepted by **CVPR 2025**.
- Created a large synthetic dataset with the GTA, generating 300,000 images across 1,200 videos with diverse driving conditions.
- Proposed a **Gaussian Mixture Model-based diffusion model** for motion planning that captures multimodal, human-like driving behavior, achieves state-of-the-art performance, accepted **by CoRL 2025**.
- Developed a **VR-integrated** simulation using **3D Gaussian Splatting** and **NeRF-based** scene reconstruction methods.
- Organized user study and collected **20,000+** diverse and feasible trajectories for benchmarking planning models in simulation.
- Captured diverse guide-dog navigation data with **XSense** and visualized it through Python-based 3D animations in **Blender**.

PKU-Agibot Lab, Peking University **Machine Learning Engineer / Research Assistant** | Jul 2022 - May 2023

- Proposed **S**tructure-**G**uided-**T**emporal-**A**ttention **Pose** as a co-first author, focusing on estimating camera-to-robot pose estimation from single-view video using temporal cross-attention, and achieved a higher precision than traditional hand-eye calibration in real time running speed (36 FPS), accepted by **CVPR 2023**.
- Created a large synthetic dataset with **Blender** containing **180,000 images** for training robust pose estimation models.
- Engineered **ROS-based** control system for **Franka Panda** robotic arm using **PyBullet** for motion planning.
- Integrated **SAM** with 6D pose estimation models, creating an end-to-end pipeline for robotic manipulation tasks on **XARM6**.

INTERN EXPERIENCE

Nanjing Zealen Technology **Machine Learning Engineer Intern** | Feb 2023 - May 2023

- Designed time-series forecasting models using ST-GCN and PyTorch for 24-hour wind power prediction.
- Implemented Temporal Fusion Transformer and XGBoost models for long-term pollution trend prediction tasks.

Beijing Siling Robot Technology **Software Development Intern** | Jan 2021 - Feb 2021

- Developed interaction interfaces and API components using C++ and QT for robotic control systems.
- Created cross-platform communication protocols for the robot operating system with low-latency performance requirements.

LEADERSHIP EXPERIENCE

Yuanpei College Woodwork class **Tutor and Course Organizer** | Mar 2020 - Jul 2023

- Founded carpentry course for engineering education, developing **hands-on curriculum** and **safety protocols** for students.
- Expanded program into co-cultivation initiative between Yuanpei College and Beijing 101 Middle School.
- Instructed **150+** students over three years, winning 2022 YuanPei Special Contribution Award scholarship.

Yuanpei College 3D Printing and Designing Lab **Founder and Leader** | Feb 2023 - Jul 2023

- Established the college's first 3D printing lab, implementing CAD-to-fabrication workflow with material printing capabilities.
- Designed graduation gifts utilizing parametric modeling techniques and additive manufacturing production methods.

SKILLS

- Programming Languages: Python, C++, C, MATLAB
- Machine Learning Frameworks: PyTorch, TensorFlow, Scikit-learn, JAX, Keras, HuggingFace
- ML Engineering: MLflow, Weights & Biases, Docker, Kubernetes, ONNX, TensorRT, Ray
- Computer Vision: OpenCV, Detectron2, YOLO, SAM, NeRF, 3D Reconstruction, Gaussian Splatting
- Data Processing: NumPy, Pandas, Matplotlib, SciPy, Dask, Spark, Luigi
- Robotics & Simulation: ROS, ROS2, Blender, PyBullet, Libfranka, Franka-Control, CAD, Fusion 360